

UNIVERSIDAD NACIONAL DE INGENIERÍA
FACULTAD DE INGENIERÍA INDUSTRIAL Y DE SISTEMAS



SI807 SISTEMAS DE INTELIGENCIA DE NEGOCIOS
“Aplicación de Sistemas de Inteligencia de Negocios en BBVA”

PRACTICA N°1

GRUPO N°4:

Espinoza Cerna, Alex	20222048D
Cárdenas Palacios Leonardo Gustavo	20222029J
Inocente Caro, Miguel Anderson	20222144C

Septiembre, 2025

Índice

- 1. ENTORNO DE LA EMPRESA
 - 1.1 Generalidades de la Empresa
 - 1.2. Identificación de Problemas de Negocio
 - 1.3. Necesidades de Información y Decisiones Críticas
 - 1.4. KPI's Iniciales
- 2. EVIDENCIA TÉCNICA
 - 2.1. Implementación de Hortonworks
 - 2.2. Diagrama de Arquitectura Inicial

Índice de Figuras

- Figura 1. Ubicación Oficina Central BBVA
- Figura 2. Organigrama BBVA
- Figura 3. Cadena de Valor BBVA
- Figura 4. Mapa de Procesos de BBVA
- Figura 5. Diagrama de Flujo de proceso de Certificación Actual BBVA
- Figura 6. Niveles de Certificación por % de Adopción del Practitioner
- Figura 7. Niveles de Certificación por % de Adopción del Continuous Integration
- Figura 8. Arquitectura Inicial

Índice de Tablas

- Tabla 1. Servicios BBVA
- Tabla 2. Descripción de funciones
- Tabla 3. Resumen de los problemas identificados
- Tabla 4. Resumen de las necesidades de información

Fuente: *Google Maps*, 2025.

D) Misión

BBVA Continental dedica gran parte de sus esfuerzos al desarrollo de la Responsabilidad Social (BBVA, 2025).

E) Visión

Búsqueda permanente de un mejor futuro para las personas, que se sostiene en tres pilares: integridad, prudencia y transparencia. (BBVA, 2025).

F) Productos

Tabla 1. Servicios BBVA

Segmento	Productos principales
Personas	Cuentas de ahorro, cuentas corrientes, depósitos a plazo, cuentas CTS, tarjetas de crédito, préstamos de consumo, hipotecarios y vehiculares.
Empresas	Créditos comerciales, factoring, leasing, comercio exterior (Comex), descuentos, emisión de deuda, avances en cuenta corriente.
Otros productos financieros	Inversiones, certificados bancarios, productos patrimoniales (Global Wealth), servicios digitales, cambio de divisas (T-Cambio).

Fuente: *Elaboración propia.*

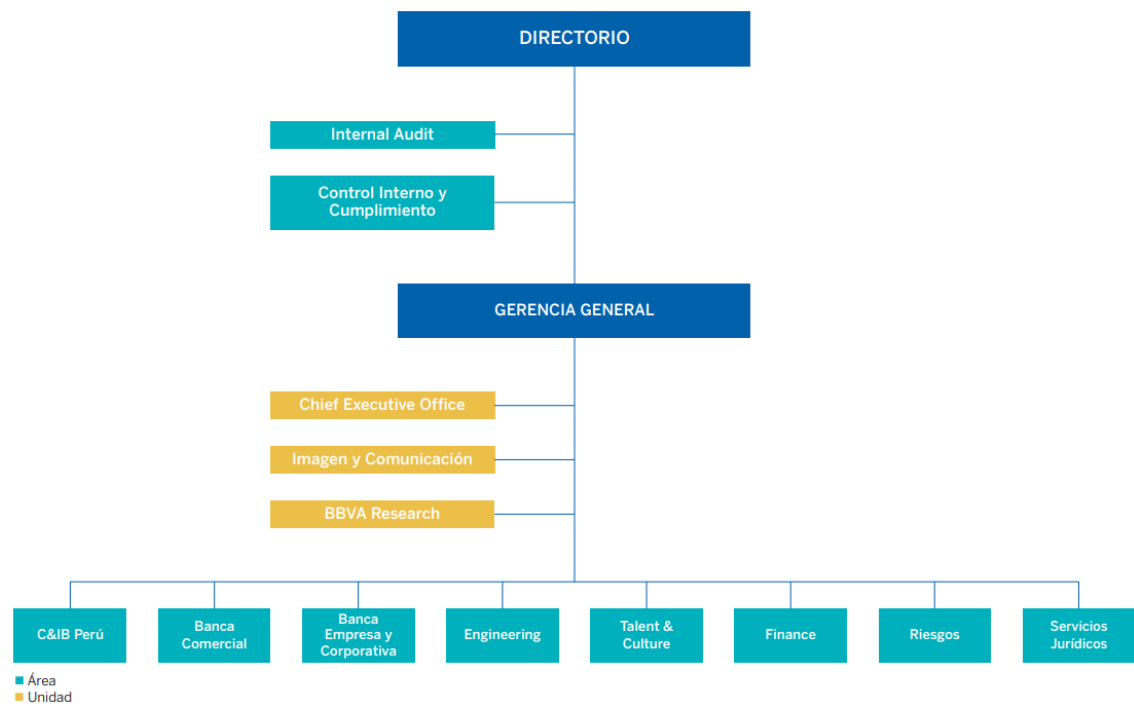
G) Clientes

Principalmente personas naturales (retail banking), pequeñas y medianas empresas (PYMEs), grandes corporaciones e instituciones. Atendidos a través de canales digitales, oficinas y banca telefónica, con enfoque en más de 80 millones de clientes globales del Grupo BBVA, adaptados al mercado peruano.

H) Organigrama

La estructura de gobierno corporativo de BBVA Perú es un componente clave para asegurar una gestión ética y sostenible, en concordancia con los objetivos estratégicos y con su compromiso de promover el desarrollo económico del país. Este esquema asegura procesos de decisión transparentes, una administración de riesgos sólida y el cumplimiento de estrictos estándares regulatorios, fortaleciendo la confianza de los grupos de interés y favoreciendo el crecimiento sostenido de la institución.

Figura 2. Organigrama BBVA



Fuente: BBVA, 2024, p. 26.

Tabla 2. Descripción de funciones

Área	Descripción
Internal Audit	Responsable de realizar evaluaciones objetivas e independientes sobre los riesgos de los procesos del negocio y los controles implementados para mitigarlos.
Control Interno y Cumplimiento	Encargada de proponer e implementar políticas relacionadas con los riesgos no financieros y de gestionar el modelo de control interno del Banco. Esta área incluye equipos como Riesgos No Financieros, Cumplimiento, Control Interno de Riesgos, Risk Control Specialists (RCS) y Relación con Supervisores.
Chief Executive Office	Responsable de la gestión de gabinete de la gerencia general. Además, apoya a todas las unidades del Banco en el alineamiento de las estrategias locales y globales para proyectos e iniciativas de impacto estratégico.
Imagen y Comunicación	Responsable de la gestión de la imagen corporativa en los diversos canales y espacios públicos.
BBVA Research	Realiza análisis económicos y proyecciones de variables clave, proporcionando información relevante para la toma de decisiones de los directivos y áreas de negocio.

Corporate & Investment Banking	Unidad especializada en banca de inversión, mercados globales, préstamos y servicios transaccionales. Ofrece un portafolio diversificado de productos financieros de alto valor para clientes corporativos e inversores institucionales.
Banca Comercial	Las unidades de Client Solutions dedicadas a Banca Retail y Banca Privada y el equipo de Red Comercial (Banca Comercial) se consolidan en una sola área para maximizar los resultados del negocio, enfocada en ofrecer soluciones financieras que respondan a las necesidades de los clientes, asegurando la sostenibilidad de las operaciones.
Banca Empresa y Corporativa	Las unidades de Client Solutions dedicadas a clientes del segmento empresas y el equipo de la Red Empresas se fusionan en una sola área que maximice la visión “end-to-end” del negocio y ofrezca soluciones financieras óptimas, liderando así la transformación del servicio del Banco y la propuesta de valor.
Engineering	Da acompañamiento a las distintas áreas de la sede central y red de oficinas para que logren sus objetivos sobre la base de soluciones y soporte tecnológico con óptimas medidas de seguridad. Asimismo, colabora en la constante mejora del diseño de procesos en ruta que promueven la productividad y el aseguramiento de actividades clave del negocio.
Talent & Culture	Sitúa al colaborador en el centro de su estrategia, lidera procesos de cambio organizacional, impulsa el fortalecimiento de la cultura organizacional y asegura espacios adecuados que potencien el bienestar y el desarrollo profesional de los equipos, al tiempo que proporciona iniciativas de diversidad e inclusión que fomenten un entorno laboral equitativo y respetuoso.
Finance	Responsable de la gestión eficiente de los recursos financieros del Banco al asegurar la relación con los inversionistas mediante la explotación y la analítica de sistemas de información.
Riesgos	Garantiza la solvencia de las operaciones del Banco mediante un modelo de gestión y evaluación de riesgos independiente y global. Esta área minimiza los impactos negativos en los resultados esperados del negocio.
Servicios Jurídicos	Asesora a las distintas unidades del Banco en temas regulatorios y legales; asimismo, diseña estrategias que aseguren la prevención y defensa para el cumplimiento de la normativa aplicable a sus actividades, a fin de evitar contingencias.

Fuente: BBVA, 2024, p. 27.

I) Cadena de Valor

BBVA organiza su cadena de valor con el objetivo de brindar servicios financieros eficientes, accesibles y digitalizados, transformando los depósitos en soluciones innovadoras que promueven la inclusión financiera y la sostenibilidad. Inspirada en el modelo de Porter, esta cadena busca optimizar procesos, potenciar el uso de la tecnología

y reforzar el compromiso social, maximizando así el valor entregado a sus clientes y a la sociedad.

A continuación, se presentan las actividades primarias y de apoyo que conforman la cadena de valor de BBVA:

Actividades Primarias:

- **Logística Interna:** Gestión de depósitos y datos financieros con cumplimiento normativo.
- **Operaciones:** Otorgamiento de créditos, transacciones y desarrollo de productos digitales (app BBVA).
- **Logística Externa:** Distribución vía oficinas, banca móvil y alianzas comerciales.
- **Marketing y Ventas:** Promoción de productos (tarjetas, préstamos) con foco en sostenibilidad.
- **Servicio Postventa:** Atención multicanal para fidelización.

Actividades de Apoyo:

- **Infraestructura:** Sistemas TI y seguridad cibernética.
- **Gestión de Recursos Humanos:** Capacitación en servicio y diversidad.
- **Desarrollo Tecnológico:** Innovación en fintech (IA, blockchain).
- **Abastecimiento:** Alianzas con proveedores tecnológicos (Google Cloud, Amazon Web Services (AWS), Red Hat, Cisco, IBM, Salesforce, etc.)

Figura 3. Cadena de Valor BBVA



Fuente: *Elaboración propia*

J) Procesos

En BBVA, los procesos constituyen la base del funcionamiento organizacional y permiten alinear la estrategia con la operación diaria. Analizar estos procesos ofrece una visión integral de cómo el banco estructura sus actividades, identifica prioridades y asegura que cada área aporte al cumplimiento de los objetivos corporativos.

Una forma de representar esta estructura es a través del mapa de procesos, que organiza las actividades en tres categorías: estratégicos, operativos y de apoyo.

a) Procesos Estratégicos: Son aquellos que marcan la dirección del banco y orientan la toma de decisiones a largo plazo. Permiten establecer lineamientos que aseguran la sostenibilidad del negocio y su capacidad de adaptación en un entorno financiero altamente competitivo y regulado. Estos son:

- Planeación estratégica corporativa: definición de metas a largo plazo, diseño de estrategias de crecimiento y adaptación a cambios del mercado.
- Gestión de riesgos y cumplimiento normativo: identificación, evaluación y mitigación de riesgos financieros, tecnológicos y regulatorios.
- Gobierno corporativo y toma de decisiones: políticas de control y estructura organizacional para garantizar transparencia y responsabilidad.
- Innovación y transformación digital: impulso de proyectos disruptivos en banca digital, IA, blockchain y fintech.

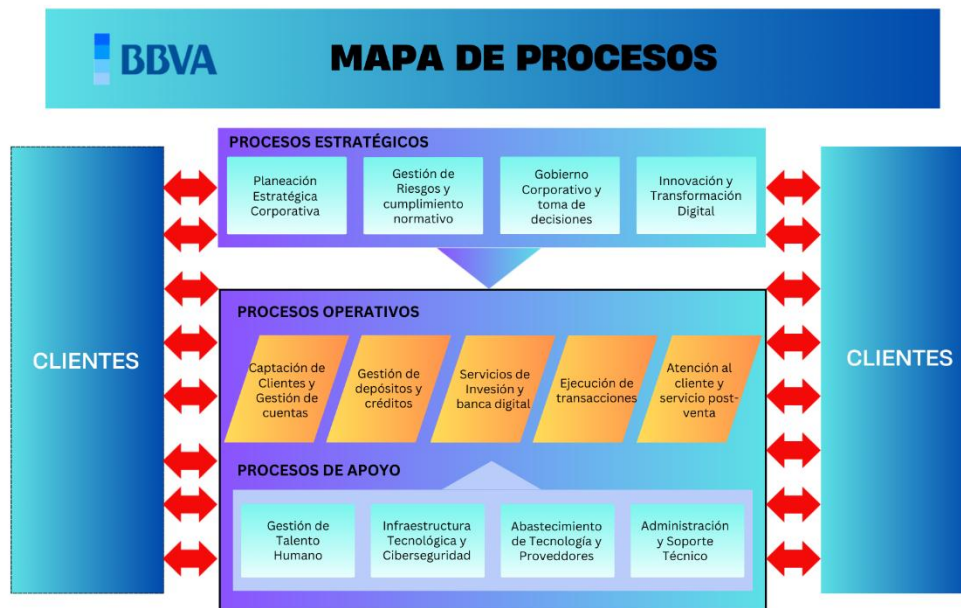
b) Procesos Operativos: Constituyen el núcleo del negocio, ya que son los que permiten a BBVA generar valor directo para los clientes. Involucran las actividades principales relacionadas con la prestación de servicios bancarios y financieros. Estos son:

- Captación de clientes y gestión de cuentas: apertura, mantenimiento y personalización de servicios financieros.
- Gestión de depósitos y créditos: administración de fondos entrantes y otorgamiento de préstamos.
- Servicios de inversión y banca digital: productos digitales, gestión de inversiones y asesoría en finanzas personales.
- Ejecución de transacciones: pagos, transferencias y operaciones financieras seguras.
- Atención al cliente y servicio postventa: soporte multicanal, resolución de incidencias y fidelización

c) Procesos de apoyo: Son los que garantizan que los procesos estratégicos y operativos puedan llevarse a cabo de manera eficiente, asegurando recursos, capacidades y soporte necesarios. Estos son:

- Gestión de talento humano: reclutamiento, capacitación y desarrollo de competencias.
- Infraestructura tecnológica y ciberseguridad: soporte de TI, plataformas digitales y protección de datos.
- Abastecimiento de tecnología y proveedores: alianzas y adquisiciones de insumos tecnológicos.
- Administración y soporte interno: finanzas internas, gestión documental y recursos logísticos.

Figura 4. Mapa de Procesos de BBVA



Fuente: *Elaboración propia*

I) Diagrama de Flujo de proceso de Certificación Actual BBVA

El proceso actual de certificación en BBVA refleja la manera en que se gestionan los datos provenientes de diferentes aplicaciones internas y cómo estos son utilizados para la toma de decisiones relacionadas con los niveles de madurez y certificación de los servicios. Este flujo corresponde a la situación As-Is, en donde todavía no se cuenta con un dashboard estandarizado, sino con el envío de datos consolidados hacia los responsables de la certificación.

El proceso inicia cuando los equipos de desarrollo generan y envían la información obtenida a través de diversas aplicaciones como Jira, Bitbucket o GIAM, las cuales registran actividades de gestión de proyectos, repositorios y accesos. Estos datos son enviados de forma no estandarizada hacia un punto central, lo que constituye la base de información con la que se trabaja actualmente.

Posteriormente, la Unidad de Analítica y BI recibe la información en bruto, la extrae y la consolida para obtener un panorama más claro de la situación de cada servicio. Este consolidado se convierte en el insumo principal que se remite al Service Owner, quien actúa como responsable de supervisar y evaluar los resultados.

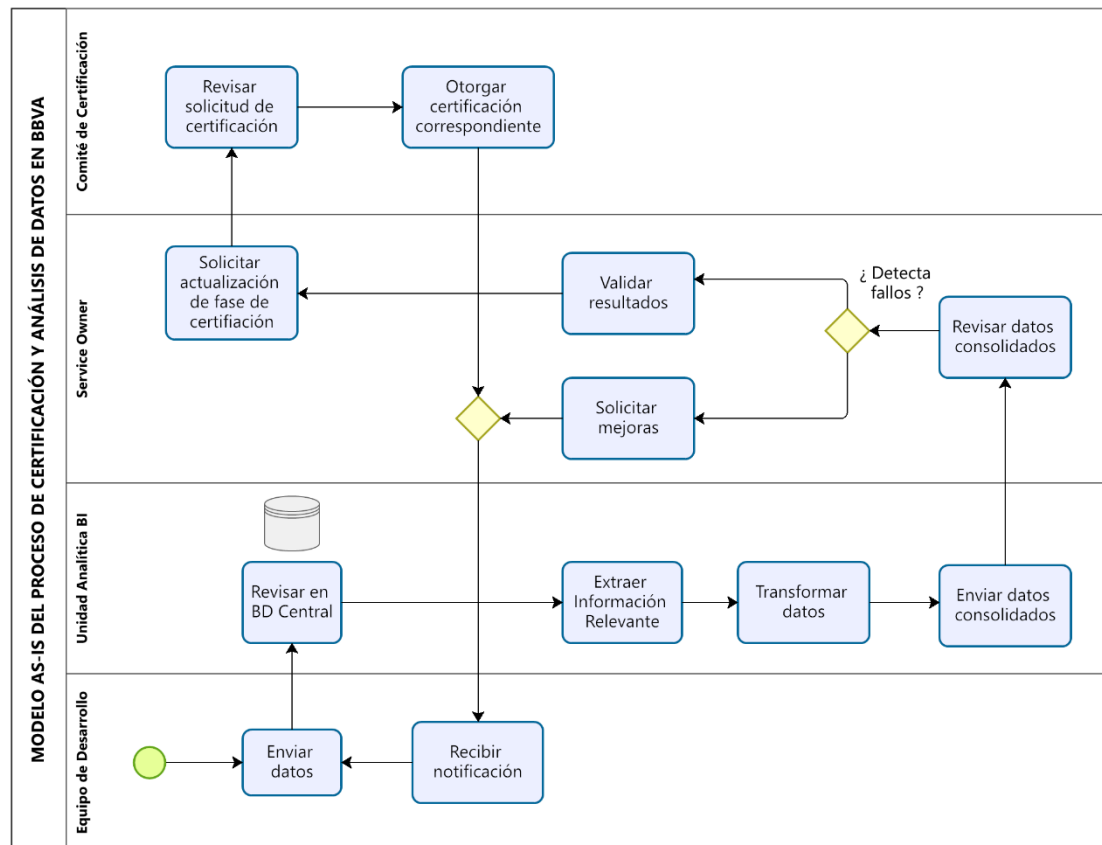
El Service Owner analiza los datos y puede tomar dos tipos de decisiones: en caso de detectar inconsistencias o deficiencias, devuelve retroalimentación al equipo de desarrollo para que se realicen las mejoras correspondientes; por otro lado, si los datos se consideran adecuados, puede solicitar la actualización del nivel de certificación del servicio (Practitioner o Continuous Integration).

Cabe resaltar que, dentro de esta última categoría, la fase Excelling cuenta con subniveles (1, 2 y 3), lo que hace más exigente la decisión de avanzar hacia una certificación superior.

Una vez que el Service Owner valida la solicitud, el caso es remitido al Comité de Certificación o área encargada de gobernanza, que revisa los criterios y otorga la certificación correspondiente. Finalmente, el equipo de desarrollo recibe la notificación formal de la certificación obtenida, cerrando así el ciclo del proceso.

Este flujo pone de manifiesto que, en la situación actual, la gestión de certificaciones en BBVA depende de datos consolidados enviados de manera manual, lo que genera limitaciones de estandarización y dificulta la trazabilidad.

Figura 5. Diagrama de Flujo de proceso de Certificación Actual BBVA



Fuente: *Elaboración propia*

1.2. Identificación de Problemas de Negocio

En el marco del modelo de madurez interna del BBVA, la medición de indicadores de adopción del Playbook del Service Owner permite identificar diversos problemas de negocio que afectan directamente la eficiencia, seguridad y calidad de los servicios tecnológicos. Estos problemas no son únicamente de carácter técnico, sino que también reflejan limitaciones en alcanzar los niveles de madurez esperados, particularmente en

Practitioner y Continuous Integration, que resultan clave para garantizar eficiencia, calidad y trazabilidad en los servicios. Dichas brechas impactan en la capacidad del banco para responder a las necesidades del mercado, asegurar la protección de la información y mantener la confianza de sus clientes. A continuación, se detallan los principales problemas de negocio detectados, junto con su descripción, implicancia y consecuencias:

A) Falta de trazabilidad y gestión de dependencias entre servicios

- **Descripción:** Los indicadores revelan la necesidad de documentar correctamente las dependencias entre servicios de distinto nivel (N1 y N2). La ausencia de esta información genera vacíos en la visibilidad de la arquitectura tecnológica y aumenta el riesgo de fallos encadenados.
- **Implicancia:** Cuando un servicio falla y no se conoce de qué otros dependen, se dificulta la identificación de la raíz del problema y la resolución de incidentes.
- **Consecuencia para el negocio:** Cuando no existe claridad sobre las dependencias entre servicios N1 y N2, una falla en un componente puede provocar interrupciones en cadena que afectan a canales digitales críticos como la banca por internet o la aplicación móvil. Esto no solo prolonga los tiempos de resolución de incidentes, sino que también eleva los costos operativos asociados a mantener la continuidad de los servicios. En consecuencia, la capacidad de respuesta frente a incidentes disminuye y el riesgo percibido por los clientes se incrementa.

B) Retrasos en la puesta de producción de servicios tecnológicos

- **Descripción:** El cumplimiento de fichas Ready For Operation (RFO) es un requisito indispensable para que los servicios puedan ser desplegados en producción. Cuando estas fichas no están completas, presentan errores o incumplen con los estándares definidos, los proyectos experimentan retrasos significativos en su liberación.
- **Implicancia:** La organización pierde capacidad de respuesta frente a competidores más ágiles, como fintechs u otros bancos digitales, lo que reduce su nivel de competitividad.
- **Consecuencia para el negocio:** El incumplimiento de las fichas RFO genera un efecto directo en los tiempos de liberación de proyectos. Esto no solo incrementa el time-to-market de las soluciones digitales, sino que además limita la capacidad del banco para aprovechar ventanas de oportunidad en el mercado. Con ello, los clientes perciben una menor agilidad en la entrega de mejoras o nuevas funcionalidades, lo que puede traducirse en insatisfacción y pérdida de competitividad frente a actores más dinámicos.

C) Baja calidad en el desarrollo de funcionalidades

- **Descripción:** El modelo considera la calidad de las *features* como un indicador clave. Cuando estas funcionalidades no cumplen con los criterios establecidos (documentación, pruebas y validaciones), se trasladan errores a producción que generan retrabajo.

- **Implicancia:** Los defectos en las funcionalidades impactan directamente en la experiencia del cliente y en la percepción de confianza hacia la plataforma digital del banco.
- **Consecuencia para el negocio:** Las funcionalidades que llegan a producción sin cumplir los criterios de calidad definidos tienden a generar errores que afectan la experiencia del cliente. Este escenario no solo provoca un aumento en reclamos y quejas, sino que también incrementa los gastos en soporte técnico y en actividades de corrección. A largo plazo, este ciclo de retrabajo desgasta a los equipos de desarrollo y reduce el tiempo que podrían destinar a la innovación, lo que limita la capacidad del banco de diferenciarse en el mercado.

D) Vulnerabilidades de seguridad en el código de los servicios

- **Descripción:** La evolución de vulnerabilidades de alto riesgo en el código es un indicador crítico para garantizar la solidez de los servicios tecnológicos. Cuando estas vulnerabilidades no son corregidas de manera oportuna, se compromete la seguridad de los datos sensibles de los clientes y la integridad de las operaciones.
- **Implicancia:** En una institución financiera, la ciberseguridad constituye un factor estratégico. La falta de control sobre vulnerabilidades expone al banco a riesgos que pueden afectar tanto su reputación como su sostenibilidad.
- **Consecuencia para el negocio:** Compromete tanto la integridad de los sistemas como la seguridad de los datos sensibles de los clientes. Si estas fallas no son detectadas y corregidas de manera oportuna, el banco queda expuesto a incidentes de fraude y a posibles fugas de información que afectarían gravemente su reputación. Además, un incidente de esta magnitud podría derivar en sanciones de organismos reguladores, lo que incrementa los costos financieros y deteriora la confianza de los usuarios en la institución.

E) Limitada visibilidad de datos integrados para la toma de decisiones

- **Descripción:** La información relativa a inventario, dependencias, calidad y seguridad de los servicios se encuentra dispersa en diferentes fuentes y plataformas. Esto obliga a los responsables a realizar procesos manuales de recopilación y análisis, lo que ralentiza la gestión de los indicadores.
- **Implicancia:** La falta de integración de datos impide contar con una visión global y en tiempo real del estado de los servicios, lo que limita la capacidad de respuesta estratégica y táctica de la organización.
- **Consecuencia para el negocio:** Obliga a los equipos a recurrir a procesos manuales de recopilación y análisis. Esto ralentiza la gestión y provoca que las decisiones se tomen con retraso o con base en datos incompletos. La consecuencia inmediata es una menor capacidad para anticipar riesgos y responder a tendencias del negocio, lo cual se traduce en ineficiencias operativas y en una pérdida de oportunidades para actuar de manera proactiva en un entorno altamente competitivo.

F) Falta de visibilidad sobre el avance en niveles de madurez

- **Descripción:** El modelo de certificación establece que los servicios deben cumplir con ciertos indicadores para avanzar del nivel Practitioner al nivel Continuous Integration. Sin embargo, la ausencia de una herramienta que consolide y muestre de forma clara el grado de cumplimiento dificulta que los responsables identifiquen brechas específicas en cada servicio.
- **Implicancia:** La organización no cuenta con un mecanismo efectivo para medir el progreso hacia niveles superiores de madurez, lo que retrasa su evolución digital y la adopción de metodologías ágiles.
- **Consecuencia para el negocio:** Genera estancamiento en niveles intermedios de madurez. Esto provoca desigualdades entre los distintos servicios, dificultando la coordinación organizacional y limitando el avance global hacia prácticas más ágiles. Como consecuencia, la institución pierde ritmo en su proceso de transformación digital y reduce su capacidad de competir con organizaciones que logran adoptar con mayor rapidez metodologías avanzadas.

Tabla 3. Resumen de los problemas identificados

Problema identificado	Categoría	Impacto principal
Retrasos en la puesta en producción de servicios tecnológicos	Eficiencia operativa / Innovación	Aumento del <i>time-to-market</i> y pérdida de oportunidades de negocio.
Falta de trazabilidad y gestión de dependencias entre servicios	Continuidad operativa / Gestión de riesgos	Riesgo de interrupciones en cascada y mayor tiempo de resolución de incidentes.
Baja calidad en el desarrollo de funcionalidades	Calidad de servicio / Experiencia del cliente	Incremento de reclamos, mayores costos de soporte y pérdida de confianza.
Vulnerabilidades de seguridad en el código de los servicios	Seguridad / Cumplimiento regulatorio	Riesgo de fraudes, fuga de información y sanciones regulatorias.
Limitada visibilidad de datos integrados para la toma de decisiones	Gestión de información / Transformación digital	Procesos de decisión más lentos y con mayor margen de error.
Falta de visibilidad sobre el avance en niveles de madurez	Madurez organizacional / Transformación digital	Estancamiento de servicios en niveles intermedios y pérdida de competitividad.

Fuente: *Elaboración propia.*

1.3. Necesidades de Información y Decisiones Críticas

Tras la identificación de los problemas de negocio, resulta necesario comprender qué tipo de decisiones deben tomarse y qué información se requiere para respaldarlas en los distintos niveles de la organización. En el caso de BBVA, la complejidad de sus procesos,

la magnitud de los datos que maneja y la presión por responder a un entorno competitivo y regulado, hacen que la información se convierta en un recurso clave para la toma de decisiones acertadas.

Las necesidades de información no son homogéneas: mientras que en el nivel estratégico se demanda una visión global y prospectiva para definir la dirección del banco, en el nivel táctico se requieren insumos que permitan coordinar, planificar y supervisar, y en el nivel operativo la información debe ser detallada y en tiempo real para resolver problemas inmediatos.

A) Nivel Estratégico: En el nivel estratégico, las decisiones se orientan a definir la dirección a largo plazo de la organización y a dar respuesta a los problemas estructurales que impactan la competitividad del banco. Retrasos en la entrega de servicios, procesos manuales y redundantes, falta de homogeneidad en la madurez de los proyectos, y riesgos en ciberseguridad son desafíos que exigen decisiones de alto alcance.

a) Tipos de decisión:

- Determinar lineamientos de transformación digital y modernización de procesos.
- Aprobar inversiones en tecnologías que reduzcan la dependencia de tareas manuales y aceleren la entrega de valor.
- Establecer políticas organizacionales en materia de gestión de riesgos y seguridad.
- Definir objetivos de largo plazo para garantizar agilidad, innovación y sostenibilidad.

b) Necesidad de información:

- Información consolidada sobre el estado general de los procesos críticos del banco.
- Análisis de riesgos y su posible impacto en clientes, regulación y reputación.
- Proyecciones de impacto financiero y operativo derivados de ineficiencias o vulnerabilidades.
- Tendencias de mercado y benchmarking con competidores que orienten la toma de decisiones estratégicas.

B) Nivel Táctico: En el nivel táctico, las decisiones buscan traducir los lineamientos estratégicos en planes concretos de gestión y coordinación, asegurando que los problemas detectados no se queden solo en un diagnóstico, sino que se conviertan en acciones sostenibles. Este nivel actúa como puente entre la visión global del banco y la ejecución diaria de los equipos.

a) Tipos de decisión:

- Asignar recursos y priorizar iniciativas en función de la relevancia de los problemas detectados.
- Diseñar planes de mejora para áreas con mayores retrasos o deficiencias en sus procesos.
- Establecer estándares comunes de trabajo que eviten redundancias y reduzcan la variabilidad entre unidades.
- Implementar mecanismos de control y seguimiento que permitan evaluar la eficacia de las acciones emprendidas.

b) Necesidad de información:

- Información segmentada por áreas, equipos o proyectos, que muestre con claridad dónde se concentran los principales problemas.
- Datos históricos de desempeño que permitan evaluar tendencias y medir el impacto de las mejoras implementadas.
- Alertas sobre riesgos operativos que puedan convertirse en bloqueos o retrasos significativos.
- Reportes de coordinación entre áreas que faciliten la identificación de dependencias críticas y posibles cuellos de botella.

C) Nivel Operativo: En el nivel operativo, las decisiones se enfocan en la ejecución diaria de tareas y procesos, con el objetivo de atender de manera inmediata los problemas que impactan la continuidad del trabajo y la calidad del servicio. Aquí es donde los equipos técnicos y funcionales ponen en práctica los lineamientos definidos a nivel táctico, asegurando que no se generen retrasos ni acumulación de fallas que escalen hacia niveles superiores.

a) Tipos de decisión:

- Organizar y priorizar actividades diarias para cumplir con plazos establecidos.
- Resolver incidencias o fallas en los sistemas en el menor tiempo posible.
- Aplicar de forma correcta los procedimientos y estándares definidos por la organización.
- Coordinar tareas de integración y despliegue para mantener la continuidad del servicio.

b) Necesidad de información:

- Datos en tiempo real sobre el estado de los procesos y tareas en ejecución.
- Alertas inmediatas de incidencias, fallas o errores críticos que requieran atención urgente.
- Registro detallado de actividades y responsables para garantizar trazabilidad en la operación.
- Información accesible y práctica que permita a los equipos ejecutar acciones rápidas y efectivas.

Tabla 4. Resumen de las necesidades de información

Nivel	Tipo de Decisión	Necesidad de Información
Estratégico	Definir lineamientos de transformación digital, aprobar inversiones clave, establecer políticas de seguridad y riesgos, y fijar objetivos de largo plazo.	Información consolidada de procesos críticos, análisis de riesgos e impacto, proyecciones financieras y operativas, y tendencias de mercado/benchmarking.
Táctico	Asignar recursos, priorizar iniciativas, diseñar planes de mejora, establecer estándares	Información segmentada por áreas y proyectos, datos históricos de desempeño, alertas de riesgos

	comunes y mecanismos de control.	operativos y reportes de coordinación entre áreas.
Operativo	Organizar actividades diarias, resolver incidencias, aplicar procedimientos definidos y coordinar despliegues.	Datos en tiempo real sobre procesos, alertas inmediatas de incidencias, registros de actividades y responsables, e información accesible para ejecutar acciones rápidas.

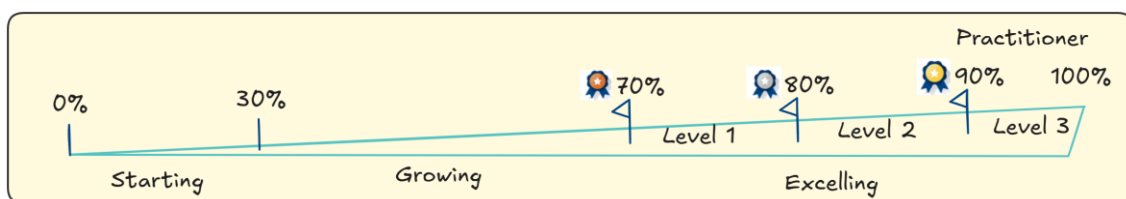
Fuente: *Elaboración propia.*

1.4. KPI's Iniciales

Para definir los KPI's se necesita definir que son los niveles de madurez Practitioner y Continuous Integration.

Practitioner: El servicio adopta prácticas básicas que permiten iniciar su camino de madurez, estableciendo una base sólida para niveles superiores.

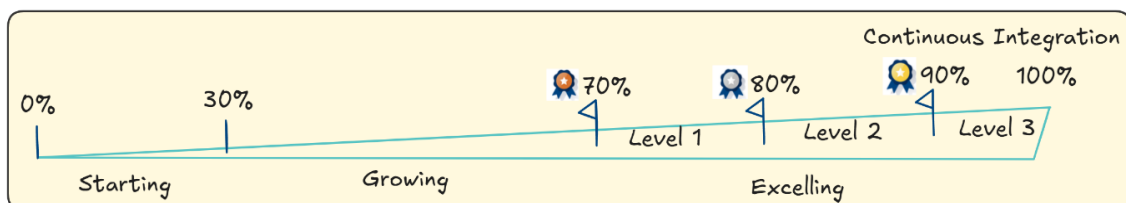
Figura 6. Niveles de Certificación por % de Adopción del Practitioner



Fuente: *Elaboración Propia.*

Continuous Integration: El servicio tiene integrado su código de manera frecuente y automática en un repositorio compartido, asegurando consistencia, detección temprana de errores y confiabilidad en el proceso de desarrollo.

Figura 7. Niveles de Certificación por % de Adopción del Continuous Integration



Fuente: *Elaboración Propia.*

Ambos niveles de madurez de estos servicios nos ayudan a tener una vista clara del SDLC (Software Development LifeCycle).

I. Indicadores para alcanzar el nivel Practitioner

A) Inventario: Agrupación en la que se miden indicadores relacionados con el inventario de servicios tecnológicos de Nucleus y de las fichas Ready For Operation (RFO) de estos servicios en Continuum.

a) Fichas de RFO para Servicios N2 de cada Servicio N1 con estados considerados “OK”

CAMPO	DESCRIPCIÓN / EJEMPLO
Nombre del KPI	Fichas de RFO para Servicios N2 de cada Servicio N1 con estados considerados “OK”
Objetivo estratégico asociado	<ul style="list-style-type: none">• Conocer / asegurar el mantenimiento del inventario tecnológico.• Definir y asegurar (cuando aplique) la estrategia de recuperación acorde con la criticidad del servicio.• Asegurar la monitorización técnica del estado del servicio.
Definición	Porcentaje de fichas de RFO para Servicios N2 de cada Servicio N1 cuyo estado se considera como “OK” para continuar subiendo a producción.
Fórmula	$\frac{\# \text{ de fichas de Servicios N2 del SN1 con RFO con status "OK"}}{\# \text{ de fichas de Servicios N2 del SN1}} \times 100\%$ <p>* Si hay una división entre cero en la fórmula, no se considera el indicador.</p>
Unidad de Medida	%
Frecuencia de Medición	Mensual
Fuente de Datos	Nucleus, Continuum
Responsable	Service Owner en el área de Engineering
Meta	≥ 90 %
Umbrales (Semáforos)	Verde: ≥ 90 %, Amarillo: 89-70%, Rojo: < 69 %
Peso del Indicador en la adopción	13% (indicador de peso bajo)

b) Servicios N2 del servicio N1 con dependencias asignadas

CAMPO	DESCRIPCIÓN / EJEMPLO
Nombre del KPI	Servicios N2 del servicio N1 con dependencias asignadas
Objetivo estratégico asociado	Conocer y asegurar el mantenimiento del inventario tecnológico del servicio y fichas RFO.
Definición	Porcentaje de Servicios N2 de cada Servicio N1 con dependencias (aquellas en las que se establece de qué otros

	servicios dependen el servicio analizado) asignadas en Nucleus.
Fórmula	$\frac{\# \text{ de Servicios N2 del Servicio N1 con dependencias}}{\# \text{ de Servicios N2 del Servicio N1}} \times 100\%$ <p>* Si hay una división entre cero en la fórmula, no se considera el indicador.</p>
Unidad de Medida	%
Frecuencia de Medición	Mensual
Fuente de Datos	Nucleus
Responsable	Service Owner en el área de Engineering
Meta	≥ 90 %
Umbrales (Semáforos)	Verde: ≥ 90 %, Amarillo: 89-70%, Rojo: < 69 %
Peso del Indicador en la adopción	13% (indicador de peso bajo)

B) Modelo Operativo: Incluye un sólo indicador que le da mayor responsabilidad al Service Owner sobre el % de Adopción de cada uno de sus Servicios N2 de los Playbooks de Desarrollo.

a) Cumplimiento del objetivo de adopción del nivel (en función de la priorización del SN2) para los Servicios N2 del Servicio N1

CAMPO	DESCRIPCIÓN / EJEMPLO
Nombre del KPI	Cumplimiento del objetivo de adopción para los Servicios N2 del Servicio N1.
Objetivo estratégico asociado	Supervisar el proceso de desarrollo de nuevas funcionalidades y participar en resolución de impedimentos.
Definición	Porcentaje de Servicios N2 de cada Servicio N1 que cumplen con el objetivo de adopción.
Fórmula	$\frac{\text{Puntaje total de adopción de SN2}}{\text{Total de SN2 medidos del SN1}} \times 100\%$ <p>* Si hay una división entre cero en la fórmula, no se considera el indicador.</p>
Unidad de Medida	%
Frecuencia de Medición	Mensual
Fuente de Datos	Chimera, Nucleus
Responsable	Service Owner en el área de Engineering
Meta	≥ 90 %
Umbrales (Semáforos)	Verde: ≥ 90 %, Amarillo: 89-70%, Rojo: < 69 %
Peso del Indicador en la adopción	27% (indicador de peso bajo)

C) Features y Desarrollos: Se cuenta con 1 indicador que permite medir la calidad de las features que se despliegan / llevan en los servicios.

a) Calidad de las Features de los Servicios N2 de cada Servicio N1

CAMPO	DESCRIPCIÓN / EJEMPLO
Nombre del KPI	Calidad de las Features de los Servicios N2 de cada Servicio N1
Objetivo estratégico asociado	<ul style="list-style-type: none"> Tener visión global del backlog de servicio y dar visibilidad a negocio. Supervisar el proceso de desarrollo de nuevas funcionalidades y participar en resolución de impedimentos.
Definición	Porcentaje de Features del SN2 de cada SN1 que cumplen con los criterios de calidad definidos.
Fórmula	$\frac{\text{Features desplegadas que cumplen con los criterios de Calidad en el } Q}{\text{Total de features desplegadas consideradas en el } Q} \times 100\%$ <p>* Si hay una división entre cero en la fórmula, no se considera el indicador.</p>
Unidad de Medida	%
Frecuencia de Medición	Mensual
Fuente de Datos	JIRA
Responsable	Service Owner en el área de Engineering
Meta	≥ 90 %
Umbrales (Semáforos)	Verde: ≥ 90 %, Amarillo: 89-70%, Rojo: < 69 %
Peso del Indicador en la adopción	20 % (indicador de peso medio)

D) Seguridad: Se ha definido un indicador cuyo objetivo es medir la seguridad del código del servicio ante las vulnerabilidades.

a) Evolución de vulnerabilidades de alto riesgo por cada mil líneas de código en las UUAA's asociadas al servicio

CAMPO	DESCRIPCIÓN / EJEMPLO
Nombre del KPI	Evolución de vulnerabilidades de alto riesgo por cada mil líneas de código en las UUAA's asociadas al servicio
Objetivo estratégico asociado	<ul style="list-style-type: none"> Garantizar el cumplimiento de regulaciones y normativas. Garantizar la seguridad del servicio ante vulnerabilidades.
Definición	Mide la variación porcentual de vulnerabilidades de alto riesgo detectadas por cada 1000 líneas de código en los repositorios asociados a un servicio (analizados en Bitbucket y mediante

	Chimera), comparando el resultado con el dato del mes anterior.
Fórmula	<ul style="list-style-type: none"> Vulnerabilidad (cada mil líneas de código): $\frac{\text{Total de vulnerabilidades high}}{\text{Total de líneas de código}} \times 1000$ Evolución de Vulnerabilidades: $\frac{\text{Vulnerabilidad del mes actual} - \text{Vulnerabilidad del mes anterior}}{\text{dato vulnerabilidad del mes anterior}} \times 100\%$ <p>* Si no se desarrolló ninguna línea de código en el mes actual, no se considera el indicador.</p> <p>* Si hay una división entre cero en la formulada de vulnerabilidades, se considera el valor de -99%.</p> <p>* Si en la evolución de vulnerabilidades sale -100% se considera como 0%.</p>
Unidad de Medida	%
Frecuencia de Medición	Mensual
Fuente de Datos	Nucleus, Quimera, BitBuckect
Responsable	Service Owner en el área de Engineering
Meta	100 %
Umbrales (Semáforos)	<ul style="list-style-type: none"> Umbral Central: Vulnerabilidad (actual) < 0.04 → 100% adopción. Vulnerabilidad (actual) > 0.2 → 0% adopción. Si el resultado se encuentra fuera del umbral central se utiliza la segunda fórmula (Evolución de vulnerabilidades) que combina la Vulnerabilidad actual y del mes anterior, y sigue la siguiente regla de umbral: Evol. Vulnerabilidad ≤ -10% → 100% adopción. Evol. Vulnerabilidad > -10% → 75% adopción. Evol. Vulnerabilidad = 0% → 50% adopción. Evol. Vulnerabilidad > 0% y Evol. Vulnerabilidad < 10% → 25 % adopción. Evol. Vulnerabilidad ≥ 10% → 0% adopción.
Peso del Indicador en la adopción	27% (indicador de peso alto)

* UUAA's: Unidades de Arquitectura y Aplicación.

E) Adopción y Certificación: Se descargan datos de distintas fuentes (Nucleus, Continuum, JIRA, Chimera) en un repositorio global. Luego se aplican las fórmulas y pesos definidos para cada indicador, redistribuyendo si alguno no aplica, y al final se obtiene el porcentaje de adopción.

a) % de adopción total del practitioner

CAMPO	DESCRIPCIÓN / EJEMPLO
Nombre del KPI	% de adopción total del practitioner
Objetivo estratégico asociado	Impulsar la adopción del nivel de practitioner en los servicios del banco para garantizar estandarización, eficiencia y madurez operativa.
Definición	Mide el porcentaje de adopción del practitioner en los servicios N1 y N2, considerando los distintos indicadores y niveles de certificación.
Fórmula	$\text{Adopción (\%)} = \frac{\{[(\text{Peso del indicador 1}) \times (\text{Resultado del indicador 1})] + [(\text{Peso del indicador 2}) \times (\text{Resultado del indicador 2})] + [(\text{Peso del indicador } \dots) \times (\text{Resultado del indicador } \dots)] + [(\text{Peso del indicador 5}) \times (\text{Resultado del indicador 5})]\}}{\{\text{Suma de los pesos de los indicadores considerados} \}}$ <p>* Si no se consideran algunos indicadores en un servicio, en vez de dividir entre 1 al Adopción (%), se dividirá entre la suma de los pesos de los indicadores que se consideran.</p>
Unidad de Medida	%
Frecuencia de Medición	Mensual
Fuente de Datos	Resultado de adopción de cada indicador.
Responsable	Service Owner en el área de Engineering
Meta	≥ 90 %
Umbrales (Semáforos)	Level 3: ≥ 90 % Level 2: 89-80 % Level 1: 79-70 % No Certificado: < 70 %

II. Indicadores para alcanzar el nivel Continuous Integration

A) Análisis y Diseño

a) % Análisis en estado ‘Analysis in Review’ menor o igual a 7 días

CAMPO	DESCRIPCIÓN / EJEMPLO
Nombre del KPI	% Análisis en estado ‘Analysis in Review’ menor o igual a 7 días
Objetivo estratégico asociado	Garantizar la finalización oportuna de los análisis manteniendo el foco en el estado “Analysis in Review”, para

	evitar retrasos y minimizar el impacto en equipos dependientes.
Definición	Porcentaje de análisis (de cualquier tipo) que han estado pendientes de revisión (con estado 'Analysis in Review') durante 7 días o menos.
Fórmula	$\frac{\text{Issues "Analysis 2.0" en 'Analysis sin Review' } \leq 7\text{días}}{\text{Total de Issues "Analysis2.0" que pasaron por 'Analysis sin Review'}} \times 100\%$ <p>* Si hay una división entre cero en la fórmula, no se considera el indicador.</p>
Unidad de Medida	%
Frecuencia de Medición	Mensual
Fuente de Datos	JIRA
Responsable	Service Owner en el área de Engineering
Meta	≥ 90 %
Umbrales (Semáforos)	Verde: ≥ 90 %, Amarillo: 89-70%, Rojo: < 69 %
Peso del Indicador en la adopción	5% (indicador de peso bajo)

B) Gestión del Backlog

a) % Historias de usuario con Release/FixVersión asociado

CAMPO	DESCRIPCIÓN / EJEMPLO
Nombre del KPI	% Historias de usuario con Release/FixVersión asociado
Objetivo estratégico asociado	Fomentar el uso de <i>releases</i> en Jira para asegurar la planificación, trazabilidad y visibilidad de los despliegues en producción, alineando los requerimientos de usuario con el código desarrollado.
Definición	Porcentaje de historias de usuario en estado " <i>Deployed</i> " durante el mes de medición que tienen el campo " <i>Fix Version</i> " informado.
Fórmula	$\frac{\text{Historias deployed con Fix Version}}{\text{Total de historias deployed}} \times 100\%$ <p>* Si hay una división entre cero en la fórmula, no se considera el indicador.</p>
Unidad de Medida	%
Frecuencia de Medición	Mensual
Fuente de Datos	JIRA
Responsable	Service Owner en el área de Engineering
Meta	≥ 90 %

Umbrales (Semáforos)	Verde: ≥ 90 %, Amarillo: 89-70%, Rojo: < 69 %
Peso del Indicador en la adopción	6% (indicador de peso bajo)

C) Desarrollo y Versionado de Código

a) % Repositorios con nomenclatura estándar

CAMPO	DESCRIPCIÓN / EJEMPLO
Nombre del KPI	% Repositorios con nomenclatura estándar
Objetivo estratégico asociado	Estandarizar la nomenclatura de ramas en los repositorios para mejorar la colaboración entre equipos, facilitar el onboarding de nuevos desarrolladores, optimizar los flujos de trabajo en CI/CD y reducir errores derivados de la falta de consistencia en los nombres.
Definición	Porcentaje de repositorios activos que cumplen con la nomenclatura establecida en las ramas. Este indicador busca verificar la adopción progresiva del estándar, inicialmente de forma opcional y posteriormente obligatorio a través de la herramienta.
Fórmula	$\frac{\text{Repositorios activos con ramas que cumplen la nomenclatura estándar}}{\text{Total de repositorios activos}} \times 100\%$ <p>* Si hay una división entre cero en la fórmula, no se considera el indicador.</p>
Unidad de Medida	%
Frecuencia de Medición	Mensual
Fuente de Datos	Bitbucket
Responsable	Service Owner en el área de Engineering
Meta	≥ 90 %
Umbrales (Semáforos)	Verde: ≥ 90 %, Amarillo: 89-70%, Rojo: < 69 %
Peso del Indicador en la adopción	6% (indicador de peso bajo)

b) Tiempo medio de aprobación de Pull Requests

CAMPO	DESCRIPCIÓN / EJEMPLO
Nombre del KPI	Tiempo medio de aprobación de Pull Requests
Objetivo estratégico asociado	Reducir los tiempos de aprobación de Pull Requests para agilizar el ciclo de desarrollo, minimizar bloqueos y fomentar buenas prácticas de revisión que mejoren la calidad del código y la eficiencia colaborativa de los equipos.

Definición	Tiempo de aprobación de Pull Requests (PR) mide cuánto tiempo pasa desde que se abre un PR hasta que es aceptada o rechazada por los revisores.
Fórmula	Mediana (tiempos en que los PR (Pull Request) han tardado en ser mergeadas o rechazadas)
Unidad de Medida	%
Frecuencia de Medición	Mensual
Fuente de Datos	Bitbucket
Responsable	Service Owner en el área de Engineering
Meta	90 %
Umbrales (Semáforos)	Tm. Aprobación < 1 día → 90% Tm. Aprobación ≥ 1 día y < 1,5 días → 70% Tm. Aprobación ≥ 1,5 días y < 3 días → 30% Tm. Aprobación ≥ 3 días → 0%
Peso del Indicador en la adopción	10% (indicador de peso medio)

c) Tamaño Medio de Pull Requests

CAMPO	DESCRIPCIÓN / EJEMPLO
Nombre del KPI	Tamaño Medio de Pull Requests
Objetivo estratégico asociado	Mantener Pull Requests pequeños facilita revisiones más rápidas y efectivas, reduce tiempos de aprobación, minimiza errores en la integración y fomenta ciclos de feedback ágiles. Esto mejora la estabilidad del código, acelera la entrega continua y refuerza la motivación del equipo al ver sus cambios aprobados e integrados con mayor rapidez.
Definición	Para fomentar integraciones pequeñas y frecuentes este indicador mide el tamaño de las integraciones (Pull Request)
Fórmula	Mediana (líneas de código modificadas en PRs mergeadas o rechazadas a ramas permanentes)
Unidad de Medida	%
Frecuencia de Medición	Mensual
Fuente de Datos	Bitbucket
Responsable	Service Owner en el área de Engineering
Meta	90 %
Umbrales (Semáforos)	Tm. PR < 300 líneas de código → 90% Tm. PR ≥ 300 y < 400 líneas de código → 70% Tm. PR ≥ 440 y < 600 líneas de código → 30% Tm. PR ≥ 600 líneas de código → 0%
Peso del Indicador en la adopción	5% (indicador de peso bajo)

d) Repositorios gobernados en el análisis estático de Seguridad

CAMPO	DESCRIPCIÓN / EJEMPLO
Nombre del KPI	Repositorios gobernados en el análisis estático de Seguridad
Objetivo estratégico asociado	Asegurar que los repositorios integrados en Chimera gestionen activamente sus vulnerabilidades, garantizando el compromiso de los equipos de desarrollo con la seguridad y minimizando el riesgo de acumulación de fallas críticas en el código.
Definición	Este indicador mide los repositorios gobernados en el análisis estático de seguridad (Chimera). Para que un repositorio esté “gobernado”, no sólo es necesario que esté integrado, sino que también los desarrolladores tengan acceso a Chimera y exista gestión activa de las debilidades de seguridad detectadas por la herramienta.
Fórmula	$\frac{\text{Repositorios activos gobernados en Chimera}}{\text{Total de repositorios activos}} \times 100\%$
Unidad de Medida	%
Frecuencia de Medición	Mensual
Fuente de Datos	Chimera, GIAM, Bitbucket
Responsable	Service Owner en el área de Engineering
Meta	$\geq 90\%$
Umbrales (Semáforos)	Verde: $\geq 90\%$, Amarillo: 89-70%, Rojo: $< 69\%$
Peso del Indicador en la adopción	7% (indicador de peso medio)

D) Análisis y construcción de código

a) Tiempo medio de integración

CAMPO	DESCRIPCIÓN / EJEMPLO
Nombre del KPI	Tiempo medio de integración
Objetivo estratégico asociado	<ul style="list-style-type: none"> Identificar cuellos de botella en revisiones, pruebas, pipelines o dependencias que ralentizan la integración. Promover buenas prácticas al reflejar la salud del pipeline y fomentar commits pequeños y revisiones rápidas. Acelerar el ciclo de desarrollo entregando valor más rápido, mejorando la calidad del software y motivando al equipo.
Definición	Tiempo desde que se hace una primera contribución a una rama hasta que se integra a rama develop, release o master/main.

Fórmula	Mediana (tiempo desde 1ª contribución → merge en ramas permanentes)
Unidad de Medida	%
Frecuencia de Medición	Mensual
Fuente de Datos	Bitbucket
Responsable	Service Owner en el área de Engineering
Meta	90 %
Umbrales (Semáforos)	Tm. Integración < 3 días → 90% Tm. Integración ≥ 3 días y < 4 días → 70% Tm. Integración ≥ 4 días y < 6 días → 30% Tm. Integración ≥ 6 días → 0%
Peso del Indicador en la adopción	16% (Indicador de peso alto)

b) Tiempo medio construcciones

CAMPO	DESCRIPCIÓN / EJEMPLO
Nombre del KPI	Tiempo medio construcciones
Objetivo estratégico asociado	<ul style="list-style-type: none"> Reducir tiempos de espera para integrar cambios fluidamente, mantener al equipo productivo y acelerar el ciclo de desarrollo. Identificar cuellos de botella en pruebas o compilaciones, fomentando buenas prácticas y mejorando la eficiencia del equipo. Detectar anomalías en los tiempos de construcción para corregir fallos rápido y asegurar un desarrollo estable y de calidad.
Definición	Mide el tiempo promedio que tardan los procesos automáticos (pipelines) en completarse. Aplica a todas las ramas y todas las compilaciones, ya que el desarrollador se ve impactado en el tiempo de ejecución de los pipelines en todas las ramas.
Fórmula	Mediana (tiempo de ejecución de pipelines correctos en cada repositorio)
Unidad de Medida	%
Frecuencia de Medición	Mensual
Fuente de Datos	Jenkins
Responsable	Service Owner en el área de Engineering
Meta	90 %
Umbrales (Semáforos)	Tm. Construcciones < 20 min → 90% Tm. Construcciones ≥ 20 min y < 35 min → 70% Tm. Construcciones ≥ 35 min y < 45 min → 30% Tm. Construcciones ≥ 45 min → 0%

Peso del Indicador en la adopción	9% (Indicador de peso medio)
-----------------------------------	------------------------------

c) % Construcciones correctas

CAMPO	DESCRIPCIÓN / EJEMPLO
Nombre del KPI	% Construcciones correctas
Objetivo estratégico asociado	<ul style="list-style-type: none"> Fomentar la responsabilidad del desarrollador aplicando buenas prácticas y asegurando calidad desde el primer commit. Contribuir a un flujo de trabajo más fluido reduciendo errores, retrocesos y demoras en la entrega. Incrementar la confianza en CI/CD, fortalecer la colaboración y promover la mejora continua del equipo.
Definición	Mide el porcentaje de pipelines que finalizan exitosamente en cualquier rama, reflejando si los cambios subidos por los desarrolladores cumplen con los criterios básicos de calidad, como la compilación correcta del código, el paso exitoso de pruebas unitarias, y la ejecución de etapas adicionales.
Fórmula	$\frac{\text{Ejecuciones de pipeline con estado SUCCESS}}{\text{Ejecuciones de pipeline con estado SUCCESS o FAILURE}} \times 100\%$ <p>* Si hay una división entre cero en la fórmula, no se considera el indicador.</p>
Unidad de Medida	%
Frecuencia de Medición	Mensual
Fuente de Datos	Jenkins
Responsable	Service Owner en el área de Engineering
Meta	≥ 90 %
Umbrales (Semáforos)	Verde: ≥ 90 %, Amarillo: 89-70%, Rojo: < 69 %
Peso del Indicador en la adopción	13% (Indicador de peso alto)

d) Tiempo medio en arreglar construcciones

CAMPO	DESCRIPCIÓN / EJEMPLO
Nombre del KPI	Tiempo medio en arreglar construcciones
Objetivo estratégico asociado	<ul style="list-style-type: none"> Promover la cultura de calidad para prevenir fallos, resolverlos rápido y fomentar la mejora continua. Reducir el tiempo de bloqueo para evitar interrupciones, efectos en cascada y retrasos en despliegues. Facilitar la detección de patrones recurrentes que permitan mejorar procesos y elevar la moral del equipo.

Definición	Mide el tiempo promedio que tarda un equipo en corregir una construcción fallida desde el momento en que ocurre el fallo hasta que se ejecuta y completa exitosamente el pipeline correspondiente.
Fórmula	Mediana (tiempo entre fallo y éxito del pipeline en ramas permanentes)
Unidad de Medida	%
Frecuencia de Medición	Mensual
Fuente de Datos	Jenkins
Responsable	Service Owner en el área de Engineering
Meta	90 %
Umbrales (Semáforos)	Tm. Arreglo < 60 min → 90% Tm. Arreglo ≥ 60 min y < 120 min → 70% Tm. Arreglo ≥ 120 min y < 180 min → 30% Tm. Arreglo ≥ 180 min → 0%
Peso del Indicador en la adopción	8% (Indicador de peso medio)

E) Testing

a) Calidad del código

CAMPO	DESCRIPCIÓN / EJEMPLO
Nombre del KPI	Calidad del código
Objetivo estratégico asociado	Asegurar que todo el código nuevo desarrollado cumple con los estándares de la industria, de manera que no se incremente la deuda técnica y se mejore progresivamente la mantenibilidad del código.
Definición	Porcentaje de repositorios con actividad en Sonar cuyas ramas develop, master, main, release y release cumplen con los criterios de calidad y cobertura de código definidos en SonarQube.
Fórmula	$\frac{\text{Repositorios con actividad y con análisis SonarQube OK}}{\text{Repositorios con actividad y con análisis SonarQube}} \times 100\%$ <p>* Si hay una división entre cero en la fórmula, no se considera el indicador.</p>
Unidad de Medida	%
Frecuencia de Medición	Mensual
Fuente de Datos	SonarQube
Responsable	Service Owner en el área de Engineering
Meta	≥ 90 %

Umbrales (Semáforos)	Verde: ≥ 90 %, Amarillo: 89-70%, Rojo: < 69 %
Peso del Indicador en la adopción	10% (Indicador de peso medio)

b) % Historias de Usuario, Dependencias y Bugs con pruebas de aceptación (XRay)

CAMPO	DESCRIPCIÓN / EJEMPLO
Nombre del KPI	% Historias de Usuario, Dependencias y Bugs con pruebas de aceptación (XRay)
Objetivo estratégico asociado	<ul style="list-style-type: none"> Mejorar el alineamiento entre las necesidades del cliente y los incrementos funcionales que el equipo desarrolla. Auditar las evidencias de ejecución de las pruebas de aceptación gracias a que se mide la cobertura funcional y su correcta ejecución.
Definición	Porcentaje de Historias de Usuario, Dependencias y Bugs de JIRA desplegados de forma planificada, que tienen resultados de ejecución de pruebas de aceptación en JIRA XRay (Test Execution con resultado final 'pass' o 'fail').
Fórmula	$\frac{\text{Ítems desplegados con pruebas Xray ejecutadas}}{\text{Ítems desplegados}} \times 100\%$ <p>* Si hay una división entre cero en la fórmula, no se considera el indicador.</p>
Unidad de Medida	%
Frecuencia de Medición	Mensual
Fuente de Datos	JIRA
Responsable	Service Owner en el área de Engineering
Meta	≥ 90 %
Umbrales (Semáforos)	Verde: ≥ 90 %, Amarillo: 89-70%, Rojo: < 69 %
Peso del Indicador en la adopción	5% (Indicador de peso bajo)

E) Adopción y Certificación:

a) % de adopción total del continuous integration

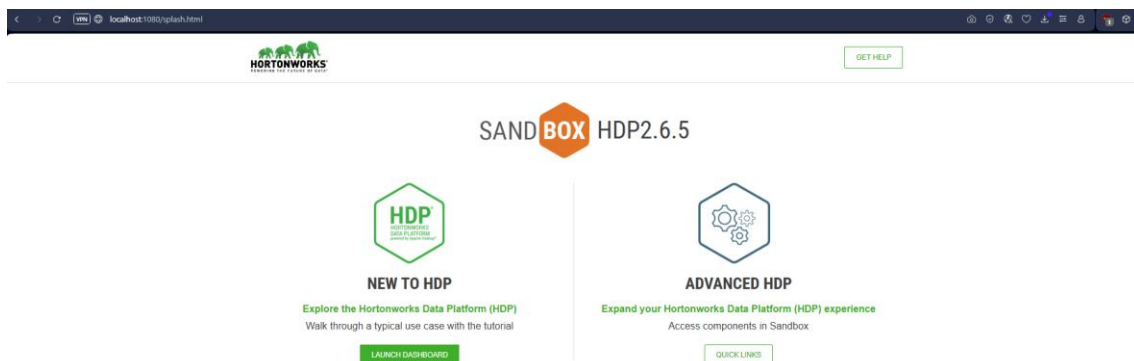
CAMPO	DESCRIPCIÓN / EJEMPLO
Nombre del KPI	% de adopción total del continuous integration
Objetivo estratégico asociado	Impulsar la adopción del nivel de continuous integration en los servicios del banco para garantizar estandarización, eficiencia y madurez operativa.

Definición	Mide el porcentaje de adopción del continuous integration en los servicios N1 y N2, considerando los distintos indicadores y niveles de certificación.
Fórmula	$\text{Adopción (\%)} = \frac{\{[(\text{Peso del indicador 1}) \times (\text{Resultado del indicador 1})] + [(\text{Peso del indicador 2}) \times (\text{Resultado del indicador 2})] + [(\text{Peso del indicador } \dots) \times (\text{Resultado del indicador } \dots)] + [(\text{Peso del indicador 12}) \times (\text{Resultado del indicador 12})]\}}{\{\text{Suma de los pesos de los indicadores considerados}\}}$ <p>* Si no se consideran algunos indicadores en un servicio, en vez de dividir entre 1 al Adopción (%), se dividirá entre la suma de los pesos de los indicadores que se consideran.</p>
Unidad de Medida	%
Frecuencia de Medición	Mensual
Fuente de Datos	Resultado de adopción de cada indicador.
Responsable	Service Owner en el área de Engineering
Meta	$\geq 90 \%$
Umbrales (Semáforos)	Level 3: $\geq 90 \%$ Level 2: 89-80% Level 1: 79-70% No Certificado: $< 70\%$

2. EVIDENCIA TÉCNICA

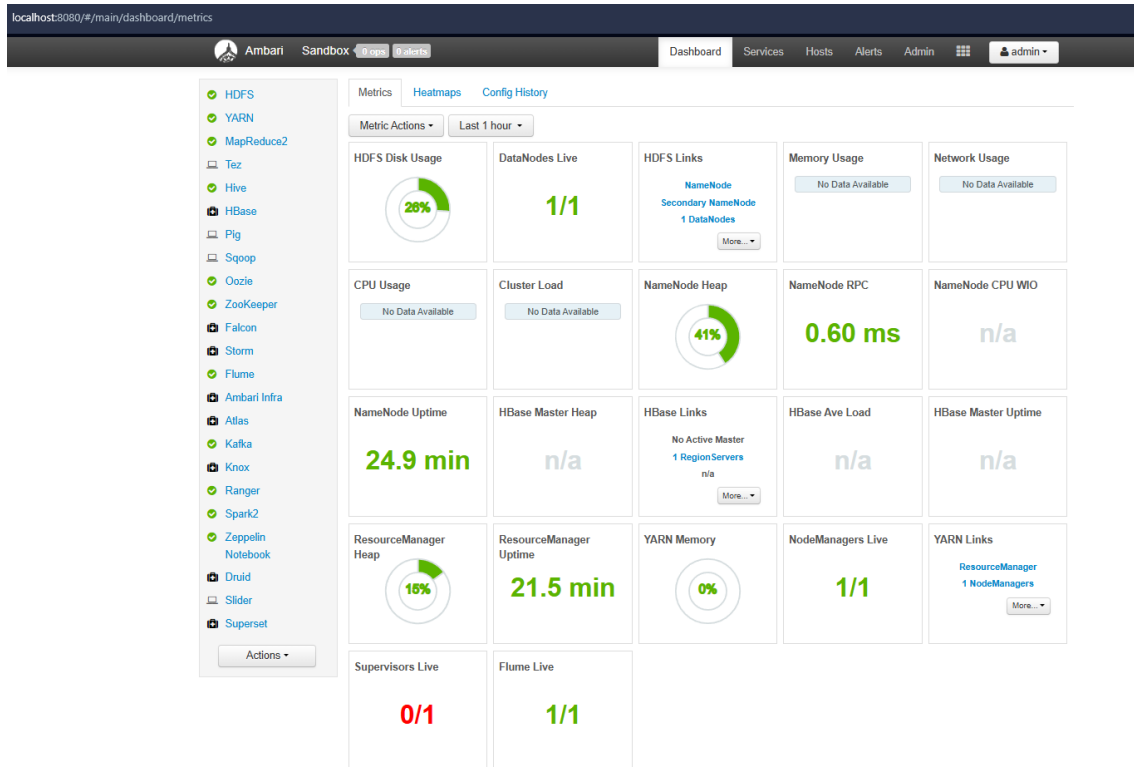
2.1. Implementación de Hortonworks

a) Entramos a Ambari desde la dirección <http://localhost:1080>.



Nota. Interfaz SandBox

b) Vemos el Dashboard de Ambari donde se visualiza los diferentes servicios en ejecución.



Nota. Dashboard con todos los servicios de Ambari

c) Modificamos la contraseña en Web Shell Client, el root pide cambiar la contraseña por primera vez y es s4ndb0x7 y luego cambiamos igual para el ambari

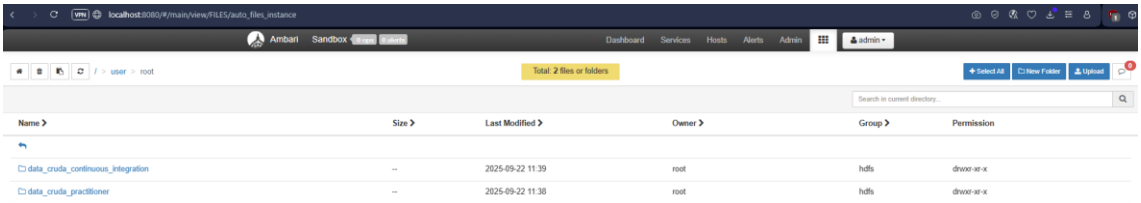
```
< > ↻ VPN localhost:4200
sandbox-hdp login: root
root@sandbox-hdp.hortonworks.com's password:
You are required to change your password immediately (root enforced)
Last login: Mon Sep 22 05:38:04 2025 from 172.18.0.2
Changing password for root.
(current) UNIX password:
New password:
Retype new password:
[root@sandbox-hdp ~]#
```

Nota. Interfaz Web Shell Client cambiando la contraseña root

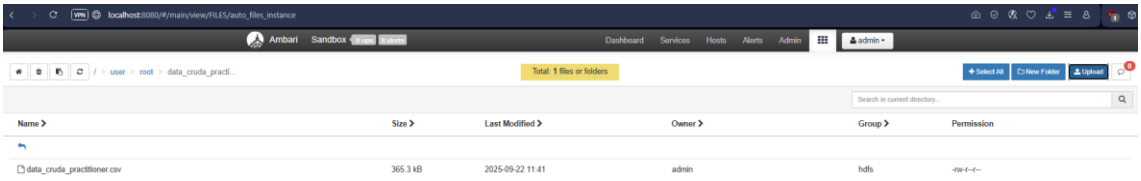
d) En HDFS colocamos los comandos para crear una carpeta y subir la data cruda relacionada al nivel de madurez Practitioner y Continuous Integration.

```
localhost:4200
sandbox-hdp login: root
root@sandbox-hdp.hortonworks.com's password:
Last login: Mon Sep 22 16:20:14 2025 from 172.18.0.2
[root@sandbox-hdp ~]# hdfs dfs -mkdir -p /user/root/data_cruda_practitioner
[root@sandbox-hdp ~]# hdfs dfs -mkdir -p /user/root/data_cruda_continuous_integration
[root@sandbox-hdp ~]#
```

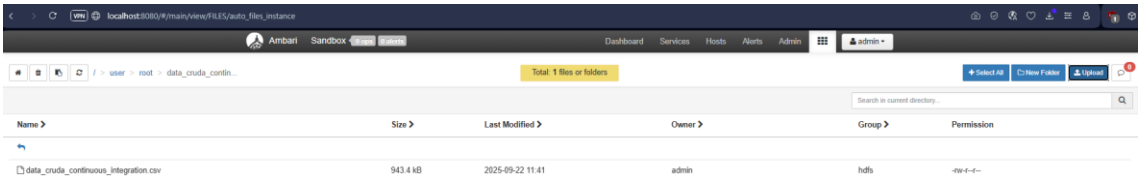
Nota. *Interfaz Web Shell Client creando las carpetas*



Nota. *Interfaz Files View con las carpetas creadas*



Nota. *Interfaz Files View con la data cruda del practitioner subida*



Nota. *Interfaz Files View con la data cruda del continuous integration subida*

e) Conectar Hive con HDFS

- Para la data cruda del Practitioner

Query Editor

Worksheet

```
1 CREATE EXTERNAL TABLE data_cruda_practitioner (  
2   mes STRING,  
3   geografia STRING,  
4   uo1_name STRING,  
5   uo12_name STRING,  
6   servicell_id INT,  
7   servicell_name STRING,  
8   nro_fichas_rfo INT,  
9   nro_sn2_ficha_rfo_status_ok INT,  
10  nro_sn2_sn1_dependencias INT,  
11  nro_sn2_sn1 INT,  
12  nro_features_desplegadas_calidad INT,  
13  nro_features_desplegadas INT,  
14  puntaje_total_adopcion_sn2 DOUBLE,  
15  total_sn2_sn1 INT,  
16  total_vulnerabilidades_high INT,  
17  total_lineas_codigo INT,  
18  total_vulnerabilidades_high_mes_ant INT,  
19  total_lineas_codigo_mes_ant INT  
20 )  
21 ROW FORMAT DELIMITED  
22 FIELDS TERMINATED BY ','  
23 STORED AS TEXTFILE  
24 LOCATION '/user/root/data_cruda_practitioner/'  
25 TBLPROPERTIES ("skip.header.line.count"="1");
```

Execute Explain Upload Save as...

New Worksheet

Nota. *Interfaz Query Editor con la query de la tabla externa del practitioner*

localhost:8090/#/main/view/HIVE/auto_hive_instance

Ambati Sandbox

Dashboard Services Hosts Alerts Admin

admin

Queries History UDFs Upload Table

Query Editor

Worksheet x Worksheet (1) x

1 Select * from data_cruda_practitioner;

Execute Explain Upload Save as...

New Worksheet

Query Process Results (Status: SUCCEEDED)

Save results...

Logs Results

Filter columns...

data_cruda_practitioner.mes	data_cruda_practitioner.geografia	data_cruda_practitioner.uo1_name	data_cruda_practitioner.uo12_name	data_cruda_practitioner.servicell_id	data_cruda_practitioner.servicell_name
03/25/2025	ARGENTINA	INGENIERIA & DATA	ALPHAARCHITECTURE	3572	ALPHA
03/25/2025	ARGENTINA	INGENIERIA & DATA	DATAHUB ARGENTINA	4343	DATIO-AR RIESGOS Y FINANZ
03/25/2025	ARGENTINA	INGENIERIA & DATA	DATAHUB ARGENTINA	4344	DATIO-AR MODELOS LOCALES
03/25/2025	ARGENTINA	INGENIERIA & DATA	DATAHUB ARGENTINA	4345	DATIO-AR MODELOS GLOBAL
03/25/2025	ARGENTINA	SYSTEMS ENGINEERING	ADQUIRENCIA	416	PA02 MERCHANTS-ACQUIRING
03/25/2025	ARGENTINA	SYSTEMS ENGINEERING	ADQUIRENCIA	4851	PA04 BBVA ADQUIRENTE - CIC
					VIDA COMERCIOS

Nota. *Interfaz Query Editor con la query para ver tabla externa del practitioner*

- Para la data cruda del Continuous Integration

```

1 CREATE EXTERNAL TABLE data_cruda_continuous_integration (
2   mes STRING,
3   geografia STRING,
4   uo11_name STRING,
5   uo12_name STRING,
6   sn1 INT,
7   sn2 INT,
8   issues_analysis2_analysis_sin_review_menor_7_dias INT,
9   total_issues_analysis2_pasaron_analysis_sin_review INT,
10  historias_deployed_fix_version INT,
11  total_historias_deployed INT,
12  repositorios_activos_nomenclatura_estandar DOUBLE,
13  total_repositorios_activos INT,
14  mediana_tiempos_pr_tardado_ser_mergeadas_o_rechazadas DOUBLE,
15  mediana_lineas_codigo_modificadas_prs_mergeadas_o_rechazadas DOUBLE,
16  repositorios_activos_gobernados_chimera INT,
17  mediana_tiempo_desde_primera_contribucion_hasta_merge_permanentes DOUBLE,
18  mediana_tiempo_ejecucion_pipeline_correctos_repositorio DOUBLE,
19  ejecuciones_pipeline_estado_success INT,
20  ejecuciones_pipeline_estado_success_o_failure INT,
21  mediana_tiempo_entre_fallo_y_exitoso_pipeline DOUBLE,
22  repositorios_actividad_y_analisis_sonarqube_ok INT,
23  repositorios_actividad_analisis_sonarqube INT,
24  items_desplegados_pruebas_xray_ejecutadas INT,
25  items_desplegados INT
26 )
27 ROW FORMAT DELIMITED
28 FIELDS TERMINATED BY ','
29 STORED AS TEXTFILE
30 LOCATION '/user/root/data_cruda_continuous_integration/'
31 TBLPROPERTIES ("skip.header.line.count"="4");
  
```

Buttons: Execute, Explain, Upload, Save as... New Worksheet

Nota. *Interfaz Query Editor con la query de la tabla externa del continuous integration*

localhost:3080/#/main/view/HIVE/auto_hive_instance

Ambari Sandbox

Queries History UDFs Upload Table

Query Editor

```
1 Select * from data_cruda_continuous_integration;
```

Buttons: Execute, Explain, Upload, Save as... New Worksheet

Query Process Results (Status: SUCCEEDED)

Save results...

Logs Results

Filter columns

data_cruda_continuous_integration.mes	data_cruda_continuous_integration.geografia	data_cruda_continuous_integration.uo11_name	data_cruda_continuous_integration.uo12_name	data_cruda_continuous_integration.sn1
06/01/2025	ESPAÑA	SYSTEMS ENGINEERING	PASIVO / PAYMENT SYSTEMS	null
06/01/2025	ARGENTINA	SYSTEMS ENGINEERING	CORPORATE FUNCTIONS	null
06/01/2025	ARGENTINA	SYSTEMS ENGINEERING	CORPORATE FUNCTIONS	null
06/01/2025	ARGENTINA	SYSTEMS ENGINEERING	CORPORATE FUNCTIONS	null
06/01/2025	ARGENTINA	SYSTEMS ENGINEERING	CORPORATE FUNCTIONS	null

previous next

Nota. *Interfaz Query Editor con la query para ver tabla externa del continuous integration*

f) Querys desde el Zeppelin

- Para la data cruda del Practitioner
 - Mostrar las 5 primeras filas, ver el tipo de dato y conteo de filas.

```
%pyspark
csv_file = "hdfs:///user/root/data_cruda_practitioner/data_cruda_practitioner.csv"
# Lee el archivo CSV usando el API de Spark
df = spark.read.option("header", "true").option("inferSchema", "true").csv(csv_file)
```

```
%pyspark
# Muestra las primeras filas de la data
df.show(5)
```

```
%pyspark
# Muestra el esquema de la data (para verificar los tipos de datos inferidos)
df.printSchema()
```

```
%pyspark
# Muestra la cantidad de registros de la data
df.count()
```

[illegible]

Nota. *Queries en Zeppelin: primeras filas, esquema y conteo de registros del practitioner*

- Cálculo de los 6 KPI's del Practitioner

```
%pyspark

from pyspark.sql import functions as F

# Cargar la tabla
df = spark.table("data_cruda_practitioner")

# --- A.a Fichas de RFO OK ---
df = df.withColumn(
    "kpi_rfo_ok",
```

```

F.when(df.nro_fichas_rfo > 0, (df.nro_sn2_ficha_rfo_status_ok / df.nro_fichas_rfo) *
100)
    .otherwise(F.lit(None))
)

# --- A.b Dependencias asignadas ---
df = df.withColumn(
    "kpi_dependencias",
    F.when(df.nro_sn2_sn1 > 0, (df.nro_sn2_sn1_dependencias / df.nro_sn2_sn1) * 100)
    .otherwise(F.lit(None))
)

# --- B.a Adopción SN2 ---
df = df.withColumn(
    "kpi_adopcion_sn2",
    F.when(df.total_sn2_sn1 > 0, (df.puntaje_total_adopcion_sn2 / df.total_sn2_sn1) *
100)
    .otherwise(F.lit(None))
)

# --- C.a Calidad Features ---
df = df.withColumn(
    "kpi_calidad_features",
    F.when(df.nro_features_desplegadas > 0, (df.nro_features_desplegadas_calidad /
df.nro_features_desplegadas) * 100)
    .otherwise(F.lit(None))
)

# --- D.a Seguridad: Vulnerabilidades ---
# Vulnerabilidad actual
df = df.withColumn(
    "vuln_actual",
    F.when(df.total_lineas_codigo > 0, (df.total_vulnerabilidades_high /
df.total_lineas_codigo) * 1000)
    .otherwise(F.lit(None))
)

# Vulnerabilidad anterior
df = df.withColumn(
    "vuln_anterior",
    F.when(df.total_lineas_codigo_mes_ant > 0, (df.total_vulnerabilidades_high_mes_ant /
df.total_lineas_codigo_mes_ant) * 1000)
    .otherwise(F.lit(None))
)

# Evolución vulnerabilidades
df = df.withColumn(
    "evol_vuln",
    F.when(df.vuln_anterior > 0, ((df.vuln_actual - df.vuln_anterior) /
df.vuln_anterior) * 100)
    .when((df.vuln_anterior == 0) & (df.vuln_actual > 0), F.lit(-99)) # regla especial
    .otherwise(F.lit(None))
)

# Adopción seguridad según reglas
df = df.withColumn(
    "kpi_seguridad",
    F.when(df.vuln_actual < 0.04, 100)
    .when(df.vuln_actual > 0.2, 0)
    .when(df.evol_vuln <= -10, 100)
    .when(df.evol_vuln == 0, 50)
    .when((df.evol_vuln > 0) & (df.evol_vuln < 10), 25)
    .when(df.evol_vuln >= 10, 0)
    .otherwise(75) # caso evol > -10 pero no cumple los otros
)

# --- E) Adopción total ---
# Pesos
peso_rfo = 0.13
peso_dependencias = 0.13
peso_adopcion_sn2 = 0.27
peso_features = 0.20

```

```

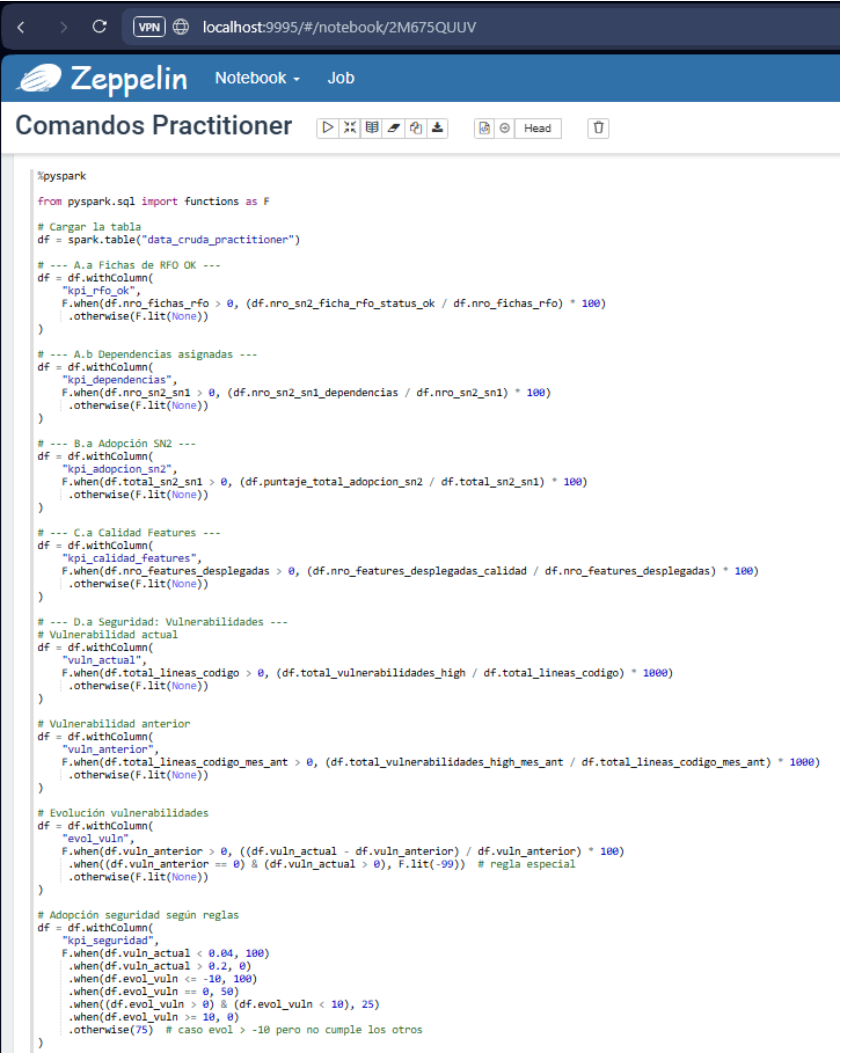
peso_seguridad = 0.27

df = df.withColumn(
    "adopcion_total",
    (
        (F.coalesce(df.kpi_rfo_ok, F.lit(0)) * peso_rfo) +
        (F.coalesce(df.kpi_dependencias, F.lit(0)) * peso_dependencias) +
        (F.coalesce(df.kpi_adopcion_sn2, F.lit(0)) * peso_adopcion_sn2) +
        (F.coalesce(df.kpi_calidad_features, F.lit(0)) * peso_features) +
        (F.coalesce(df.kpi_seguridad, F.lit(0)) * peso_seguridad)
    ) / (
        (F.when(df.kpi_rfo_ok.isNotNull(), peso_rfo).otherwise(0)) +
        (F.when(df.kpi_dependencias.isNotNull(), peso_dependencias).otherwise(0)) +
        (F.when(df.kpi_adopcion_sn2.isNotNull(), peso_adopcion_sn2).otherwise(0)) +
        (F.when(df.kpi_calidad_features.isNotNull(), peso_features).otherwise(0)) +
        (F.when(df.kpi_seguridad.isNotNull(), peso_seguridad).otherwise(0))
    )
)

# Mostrar resultados
df_filtered = df.filter(df.mes != "mes")

df_filtered.select(
    "mes", "geografia", "uol1_name", "uol2_name",
    "servicell_id", "servicell_name",
    "kpi_rfo_ok", "kpi_dependencias", "kpi_adopcion_sn2",
    "kpi_calidad_features", "kpi_seguridad", "adopcion_total"
).show(20, False)

```



The screenshot shows a Zeppelin Notebook interface with a dark theme. The browser address bar indicates the URL is localhost:9995/#/notebook/2M675QUUV. The notebook title is "Comandos Practitioner". The code is written in PySpark and includes comments in Spanish. It defines several metrics: kpi_rfo_ok, kpi_dependencias, kpi_adopcion_sn2, kpi_calidad_features, kpi_seguridad, and evol_vuln. The script uses F.when and F.coalesce functions to handle null values and calculate weighted averages. The final output is shown as a table with columns for mes, geografia, uol1_name, uol2_name, servicell_id, servicell_name, kpi_rfo_ok, kpi_dependencias, kpi_adopcion_sn2, kpi_calidad_features, kpi_seguridad, and adopcion_total.

```

%pyspark

from pyspark.sql import functions as F

# Cargar la tabla
df = spark.table("data_cruda_practitioner")

# --- A.a Fichas de RFO OK ---
df = df.withColumn(
    "kpi_rfo_ok",
    F.when(df.nro_sn2_sn1 > 0, (df.nro_sn2_sn1_fichas_rfo_status_ok / df.nro_fichas_rfo) * 100)
    .otherwise(F.lit(None))
)

# --- A.b Dependencias asignadas ---
df = df.withColumn(
    "kpi_dependencias",
    F.when(df.nro_sn2_sn1 > 0, (df.nro_sn2_sn1_dependencias / df.nro_sn2_sn1) * 100)
    .otherwise(F.lit(None))
)

# --- B.a Adopción SN2 ---
df = df.withColumn(
    "kpi_adopcion_sn2",
    F.when(df.total_sn2_sn1 > 0, (df.puntaje_total_adopcion_sn2 / df.total_sn2_sn1) * 100)
    .otherwise(F.lit(None))
)

# --- C.a Calidad Features ---
df = df.withColumn(
    "kpi_calidad_features",
    F.when(df.nro_features_desplegadas > 0, (df.nro_features_desplegadas_calidad / df.nro_features_desplegadas) * 100)
    .otherwise(F.lit(None))
)

# --- D.a Seguridad: Vulnerabilidades ---
# Vulnerabilidad actual
df = df.withColumn(
    "vuln_actual",
    F.when(df.total_lineas_codigo > 0, (df.total_vulnerabilidades_high / df.total_lineas_codigo) * 1000)
    .otherwise(F.lit(None))
)

# Vulnerabilidad anterior
df = df.withColumn(
    "vuln_anterior",
    F.when(df.total_lineas_codigo_mes_ant > 0, (df.total_vulnerabilidades_high_mes_ant / df.total_lineas_codigo_mes_ant) * 1000)
    .otherwise(F.lit(None))
)

# Evolución vulnerabilidades
df = df.withColumn(
    "evol_vuln",
    F.when(df.vuln_anterior > 0, ((df.vuln_actual - df.vuln_anterior) / df.vuln_anterior) * 100)
    .when(df.vuln_anterior == 0 & (df.vuln_actual > 0), F.lit(-99)) # regla especial
    .otherwise(F.lit(None))
)

# Adopción seguridad según reglas
df = df.withColumn(
    "kpi_seguridad",
    F.when(df.vuln_actual < 0.04, 100)
    .when(df.vuln_actual > 0.2, 0)
    .when(df.evol_vuln <= -10, 100)
    .when(df.evol_vuln == 0, 50)
    .when((df.evol_vuln > 0) & (df.evol_vuln < 10), 25)
    .when(df.evol_vuln >= 10, 0)
    .otherwise(75) # caso evol > -10 pero no cumple los otros
)

```

```
# --- E) Adopción total ---
# Pesos
peso_rfo = 0.13
peso_dependencias = 0.13
peso_adopcion_sn2 = 0.27
peso_features = 0.20
peso_seguridad = 0.27

df = df.withColumn(
    "adopcion_total",
    (
        (F.coalesce(df.kpi_rfo_ok, F.lit(0)) * peso_rfo) +
        (F.coalesce(df.kpi_dependencias, F.lit(0)) * peso_dependencias) +
        (F.coalesce(df.kpi_adopcion_sn2, F.lit(0)) * peso_adopcion_sn2) +
        (F.coalesce(df.kpi_calidad_features, F.lit(0)) * peso_features) +
        (F.coalesce(df.kpi_seguridad, F.lit(0)) * peso_seguridad)
    ) / (
        (F.when(df.kpi_rfo_ok.isNotNull(), peso_rfo).otherwise(0)) +
        (F.when(df.kpi_dependencias.isNotNull(), peso_dependencias).otherwise(0)) +
        (F.when(df.kpi_adopcion_sn2.isNotNull(), peso_adopcion_sn2).otherwise(0)) +
        (F.when(df.kpi_calidad_features.isNotNull(), peso_features).otherwise(0)) +
        (F.when(df.kpi_seguridad.isNotNull(), peso_seguridad).otherwise(0))
    )
)

# Mostrar resultados
df_filtered = df.filter(df.mes != "mes")

df_filtered.select(
    "mes", "geografia", "uol1_name", "uol2_name",
    "servicell_id", "servicell_name",
    "kpi_rfo_ok", "kpi_dependencias", "kpi_adopcion_sn2",
    "kpi_calidad_features", "kpi_seguridad", "adopcion_total"
).show(20, False)
```

mes	geografia	uol1_name	uol2_name	servicell_id	servicell_name	kpi_rfo_ok	kpi_dependencias	kpi_adopcion_sn2	kpi_calidad_features	kpi_seguridad	adopcion_total
03/25/2025	ARGENTINA	INGENIERIA & DATA	ALPHA ARCHITECTURE	3572	ALPHA	100.0	100.0	100.0	98.76543209876543	75	93.00308641975309
03/25/2025	ARGENTINA	INGENIERIA & DATA	DATAHUB ARGENTINA	4343	DATAID-AR RIESGOS Y FINANZAS	100.0	100.0	100.0	86.0	75	62.25
03/25/2025	ARGENTINA	INGENIERIA & DATA	DATAHUB ARGENTINA	4344	DATAID-AR MODELOS LOCALES	100.0	100.0	100.0	100.0	75	93.25
03/25/2025	ARGENTINA	INGENIERIA & DATA	DATAHUB ARGENTINA	4345	DATAID-AR MODELOS GLOBALES	100.0	100.0	100.0	100.0	75	93.25
03/25/2025	ARGENTINA	SYSTEMS ENGINEERING	ADQUIRENCIA	416	PA02 MERCHANTS-ACQUIRING / COMERCIOS-ADQUIRENCIA	100.0	100.0	100.0	100.0	75	79.75
03/25/2025	ARGENTINA	SYSTEMS ENGINEERING	ADQUIRENCIA	4851	PA04 BBVA ADQUIRENTE - CICLO DE VIDA COMERCIOS	100.0	100.0	100.0	100.0	75	66.25
03/25/2025	ARGENTINA	SYSTEMS ENGINEERING	ADQUIRENCIA	4853	PA03 PSP - MERCURIO	100.0	100.0	100.0	100.0	75	66.25
03/25/2025	ARGENTINA	SYSTEMS ENGINEERING	ADQUIRENCIA	5297	PA05 BBVA ADQUIRENTE - TRANSACCIONALIDAD	100.0	100.0	100.0	100.0	75	79.75
03/25/2025	ARGENTINA	SYSTEMS ENGINEERING	CANALES EMPRESAS	379	AR02 ENTERPRISE WEB CORE	50.0	100.0	100.0	100.0	75	75.0
03/25/2025	ARGENTINA	SYSTEMS ENGINEERING	CANALES EMPRESAS	381	EB02 ENTERPRISE MOBILE / MOVILES EMPRESAS	100.0	0.0	100.0	100.0	75	62.7358495668377
03/25/2025	ARGENTINA	SYSTEMS ENGINEERING	CANALES EMPRESAS	382	EB01 ENTERPRISE WEB / WEB EMPRESAS	100.0	25.0	100.0	100.0	75	56.5
03/25/2025	ARGENTINA	SYSTEMS ENGINEERING	CANALES EMPRESAS	4264	AR37 WEB PUBLICA EMPRESAS	50.0	0.0	100.0	100.0	75	58.471698113207545
03/25/2025	ARGENTINA	SYSTEMS ENGINEERING	CANALES EMPRESAS	4651	EB03 CANALES EMPRESAS	0.0	0.0	100.0	100.0	75	65.5625
03/25/2025	ARGENTINA	SYSTEMS ENGINEERING	CANALES EMPRESAS	5963	AR03 OPEN BANKING EMPRESAS	100.0	0.0	100.0	100.0	75	66.273972602739725

Nota. Cálculo de los 6 KPI's del Practitioner en Zeppelin.

- Para la data cruda del Continuous Integration

➤ Mostrar las 5 primeras filas, ver el tipo de dato y conteo de filas.

```
%pyspark
csv_file =
"hdfs://user/root/data_cruda_continuous_integration/data_cruda_continuous_integration.csv"
# Lee el archivo CSV usando el API de Spark
df = spark.read.option("header", "true").option("inferSchema", "true").csv(csv_file)

%pyspark
# Muestra las primeras filas de la data
df_subset = df.limit(8).subtract(df.limit(3))
df_subset.show()

%pyspark
# Muestra el esquema de la data (para verificar los tipos de datos inferidos)
df.printSchema()

%pyspark
# Muestra la cantidad de registros de la data
df.count()
```

Comandos Continuous Integration

Command 1: Spark SQL query to select the first 10 rows of the 'data_cruda_continuous_integration' table.

```
%spark
# Selecciona las primeras 10 filas de la data
df_subset = df.limit(10).subtract(df.limit(1))
df_subset.show()
```

Command 2: Spark SQL query to calculate 13 KPIs for the same table.

```
%spark
-- mes: string (nullable = true)
-- geoip: string (nullable = true)
-- url_name: string (nullable = true)
-- url_code: string (nullable = true)
-- url: string (nullable = true)
-- url: string (nullable = true)
-- issues_analysis2_analysis_sin_review_menor_7_dias: integer (nullable = true)
-- total_issues_analysis2_pasaron_analysis_sin_review: integer (nullable = true)
-- historias_deployed_fix_version: integer (nullable = true)
-- total_historias_deployed: integer (nullable = true)
-- repository_active_nomenclatura_estandar: integer (nullable = true)
-- total_repository_active: integer (nullable = true)
-- mediana_tiempo_pr_tardado_ser_mergeadas_o_rechazadas: double (nullable = true)
-- mediana_lineas_codigo_modificadas_prs_mergeadas_o_rechazadas: double (nullable = true)
-- repository_active_governados_chimera: integer (nullable = true)
-- mediana_tiempo_desde_primera_contribucion_hasta_merge: double (nullable = true)
```

Command 3: Spark SQL query to count the number of records in the table.

```
%spark
# Cuenta la cantidad de registros de la data
df.count()
```

Nota. Queries en Zeppelin: primeras filas, esquema y conteo de registros del continuous integration

➤ Cálculo de los 13 KPI's del Continuous Integration

```
%pyspark
```

```
from pyspark.sql import functions as F
```

```
# 1) Cargar la tabla
```

```
df = spark.table("data_cruda_continuous_integration")
```

```
# 2) Quitar filas basura (headers/nulos en 'mes' + eliminar las 3 primeras filas "permanentes", "repositorio", "mediana tiempo...")
```

```
df = df.filter((F.col("mes").isNotNull()) & (~F.col("mes").rlike("(?)^mes$")))
df = df.filter(~F.col("mes").rlike("(?)^permanentes|^repositorio|^mediana_tiempo"))
```

```
# 3) Convertir columnas a double
```

```
to_double = [
    "issues_analysis2_analysis_sin_review_menor_7_dias",
    "total_issues_analysis2_pasaron_analysis_sin_review",
    "historias_deployed_fix_version",
    "total_historias_deployed",
    "repositorios_activos_nomenclatura_estandar",
    "total_repositorios_activos",
    "mediana_tiempo_pr_tardado_ser_mergeadas_o_rechazadas",
    "mediana_lineas_codigo_modificadas_prs_mergeadas_o_rechazadas",
    "repositorios_activos_governados_chimera",
    "mediana_tiempo_desde_primera_contribucion_hasta_merge_permanentes",
    "mediana_tiempo_ejecucion_pipeline_correctos_repositorio",
    "ejecuciones_pipeline_estado_success",
    "ejecuciones_pipeline_estado_success_o_failure",
    "mediana_tiempo_entre_fallo_y_exito_pipeline",
    "repositorios_actividad_y_analisis_sonarqube_ok",
    "repositorios_actividad_analisis_sonarqube",
    "items_desplegados_pruebas_xray_ejecutadas",
    "items_desplegados"
]
```

```
for c in to_double:
```

```
    if c in df.columns:
```

```
        df = df.withColumn(c, F.col(c).cast("double"))
```

```
# 3) Helper para escalar medianas en % (min = 100%, max = 0%)
```



```

def linear_scale_col(col, min_val, max_val):
    return (
        F.when(col.isNull(), None)
        .when(col <= F.lit(min_val), F.lit(100.0))
        .when(col >= F.lit(max_val), F.lit(0.0))
        .otherwise(100.0 - ((col - F.lit(min_val)) * (100.0 / (F.lit(max_val) -
F.lit(min_val)))))
    )

# 4) Cálculo de KPIs en %
df_kpis = df \
    .withColumn("kpi_analysis_le_7d",
        F.when(F.col("total_issues_analysis2_pasaron_analysis_sin_review") > 0,
            (F.col("issues_analysis2_analysis_sin_review_menor_7_dias") /
            F.col("total_issues_analysis2_pasaron_analysis_sin_review")) * 100)) \
    .withColumn("kpi_fixversion",
        F.when(F.col("total_historias_deployed") > 0,
            (F.col("historias_deployed_fix_version") /
            F.col("total_historias_deployed")) * 100)) \
    .withColumn("kpi_nomenclatura",
        F.when(F.col("total_repositorios_activos") > 0,
            (F.col("repositorios_activos_nomenclatura_estandar") /
            F.col("total_repositorios_activos")) * 100)) \
    .withColumn("kpi_tiempo_aprobacion_pr",
        linear_scale_col(F.col("mediana_tiempos_pr_tardado_ser_mergeadas_o_rechazadas"),
24, 72)) \
    .withColumn("kpi_tamano_pr",

linear_scale_col(F.col("mediana_lineas_codigo_modificadas_prs_mergeadas_o_rechazadas"),
400, 1200)) \
    .withColumn("kpi_chimera",
        F.when(F.col("total_repositorios_activos") > 0,
            (F.col("repositorios_activos_gobernados_chimera") /
            F.col("total_repositorios_activos")) * 100)) \
    .withColumn("kpi_tiempo_integracion",

linear_scale_col(F.col("mediana_tiempo_desde_primera_contribucion_hasta_merge_permanente
s"), 1, 5)) \
    .withColumn("kpi_tiempo_construccion",

linear_scale_col(F.col("mediana_tiempo_ejecucion_pipelines_correctos_repositorio"), 10,
60)) \
    .withColumn("kpi_build_success",
        F.when(F.col("ejecuciones_pipeline_estado_success_o_failure") > 0,
            (F.col("ejecuciones_pipeline_estado_success") /
            F.col("ejecuciones_pipeline_estado_success_o_failure")) * 100)) \
    .withColumn("kpi_tiempo_arreglo",
        linear_scale_col(F.col("mediana_tiempo_entre_fallo_y_exitito_pipeline"), 1, 8)) \
    .withColumn("kpi_calidad_codigo",
        F.when(F.col("repositorios_actividad_analisis_sonarqube") > 0,
            (F.col("repositorios_actividad_y_analisis_sonarqube_ok") /
            F.col("repositorios_actividad_analisis_sonarqube")) * 100)) \
    .withColumn("kpi_xray",
        F.when(F.col("items_desplegados") > 0,
            (F.col("items_desplegados_pruebas_xray_ejecutadas") /
            F.col("items_desplegados")) * 100))

# 5) Pesos
weights = {
    "kpi_analysis_le_7d": 0.05,
    "kpi_fixversion": 0.06,
    "kpi_nomenclatura": 0.06,
    "kpi_tiempo_aprobacion_pr": 0.10,
    "kpi_tamano_pr": 0.05,
    "kpi_chimera": 0.07,
    "kpi_tiempo_integracion": 0.16,
    "kpi_tiempo_construccion": 0.09,
    "kpi_build_success": 0.13,
    "kpi_tiempo_arreglo": 0.08,
    "kpi_calidad_codigo": 0.10,
    "kpi_xray": 0.05
}

```

```

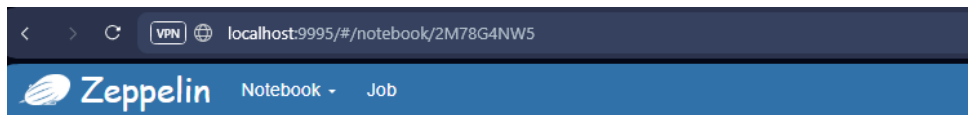
# 6) Cálculo de adopción total ponderada
num_expr = None
den_expr = None
for colname, w in weights.items():
    expr_num = F.when(F.col(colname).isNotNull(), F.col(colname) *
F.lit(w)).otherwise(F.lit(0.0))
    expr_den = F.when(F.col(colname).isNotNull(), F.lit(w)).otherwise(F.lit(0.0))
    num_expr = expr_num if num_expr is None else (num_expr + expr_num)
    den_expr = expr_den if den_expr is None else (den_expr + expr_den)

df_kpis = df_kpis \
    .withColumn("adopcion_total_num", num_expr) \
    .withColumn("adopcion_total_den", den_expr) \
    .withColumn("adopcion_total",
        F.when(F.col("adopcion_total_den") > 0,
            (F.col("adopcion_total_num") / F.col("adopcion_total_den"))
        ).otherwise(F.lit(None)))

# 7) Selección final y mostrar
cols_show = [
    "mes", "geografia", "uol1_name", "uol2_name",
    "kpi_analysis_le_7d", "kpi_fixversion", "kpi_nomenclatura",
    "kpi_tiempo_aprobacion_pr", "kpi_tamano_pr", "kpi_chimera",
    "kpi_tiempo_integracion", "kpi_tiempo_construccion",
    "kpi_build_success", "kpi_tiempo_arreglo",
    "kpi_calidad_codigo", "kpi_xray", "adopcion_total"
]

existing_cols_show = [c for c in cols_show if c in df_kpis.columns]
df_kpis.select(existing_cols_show).show(50, False)

```



Comandos Continuous Integration

```
%pyspark
from pyspark.sql import functions as F

# 1) Cargar la tabla
df = spark.table("data_cruda_continuous_integration")

# 2) Quitar filas basura (headers/nulos en 'mes' + eliminar las 3 primeras filas "permanentes", "repositorio", "mediana_tiempo...")
df = df.filter((F.col("mes").isNull()) & (~F.col("mes").rlike("(?)mes$")))
df = df.filter(~F.col("mes").rlike("(?)permanentes|^repositorio|^mediana_tiempo"))

# 3) Convertir columnas a double
to_double = [
    "issues_analysis2_analysis_sin_review_menor_7_dias",
    "total_issues_analysis2_pasaron_analysis_sin_review",
    "historias_deployed_fix_version",
    "total_historias_deployed",
    "repositorios_activos_nomenclatura_estandar",
    "total_repositorios_activos",
    "mediana_tiempos_pr_tardado_ser_mergeadas_o_rechazadas",
    "mediana_lineas_codigo_modificadas_prs_mergeadas_o_rechazadas",
    "repositorios_activos_gobernados_chimera",
    "mediana_tiempo_desde_primera_contribucion_hasta_merge_permanentes",
    "mediana_tiempo_ejecucion_pipeline_correctos_repositorio",
    "ejecuciones_pipeline_estado_success",
    "ejecuciones_pipeline_estado_success_o_failure",
    "mediana_tiempo_entre_fallo_y_exitos_pipeline",
    "repositorios_actividad_y_analisis_sonarqube_ok",
    "repositorios_actividad_analisis_sonarqube",
    "items_desplegados_pruebas_xray_ejecutadas",
    "items_desplegados"
]

for c in to_double:
    if c in df.columns:
        df = df.withColumn(c, F.col(c).cast("double"))

# 3) Helper para escalar medianas en % (min = 100%, max = 0%)
def linear_scale_col(col, min_val, max_val):
    return (
        F.when(col.isNull(), None)
        .when(col <= F.lit(min_val), F.lit(100.0))
        .when(col >= F.lit(max_val), F.lit(0.0))
        .otherwise(100.0 - ((col - F.lit(min_val)) * (100.0 / (F.lit(max_val) - F.lit(min_val)))))
    )

# 4) Cálculo de KPIs en %
df_kpis = df \
    .withColumn("kpi_analysis_le_7d",
        F.when(F.col("total_issues_analysis2_pasaron_analysis_sin_review") > 0,
            (F.col("issues_analysis2_analysis_sin_review_menor_7_dias") /
             F.col("total_issues_analysis2_pasaron_analysis_sin_review") * 100)) \
    .withColumn("kpi_fixversion",
        F.when(F.col("total_historias_deployed") > 0,
            (F.col("historias_deployed_fix_version") /
             F.col("total_historias_deployed") * 100)) \
    .withColumn("kpi_nomenclatura",
        F.when(F.col("total_repositorios_activos") > 0,
            (F.col("repositorios_activos_nomenclatura_estandar") /
             F.col("total_repositorios_activos") * 100)) \
    .withColumn("kpi_tiempo_aprobacion_pr",
        linear_scale_col(F.col("mediana_tiempos_pr_tardado_ser_mergeadas_o_rechazadas"), 24, 72)) \
    .withColumn("kpi_tamano_pr",
        linear_scale_col(F.col("mediana_lineas_codigo_modificadas_prs_mergeadas_o_rechazadas"), 400, 1200)) \
    .withColumn("kpi_chimera",

F.when(F.col("total_repositorios_activos") > 0,
        (F.col("repositorios_activos_gobernados_chimera") /
         F.col("total_repositorios_activos") * 100)) \
    .withColumn("kpi_tiempo_ejecucion_pipeline_correctos_repositorio",
        F.when(F.col("ejecuciones_pipeline_estado_success") > 0,
            (F.col("ejecuciones_pipeline_estado_success") /
             F.col("ejecuciones_pipeline_estado_success_o_failure") * 100)) \
    .withColumn("kpi_tiempo_entre_fallo_y_exitos_pipeline",
        F.when(F.col("mediana_tiempo_entre_fallo_y_exitos_pipeline") > 0,
            (F.col("mediana_tiempo_entre_fallo_y_exitos_pipeline") /
             F.col("mediana_tiempo_entre_fallo_y_exitos_pipeline") * 100)) \
    .withColumn("kpi_actividad_y_analisis_sonarqube_ok",
        F.when(F.col("repositorios_actividad_y_analisis_sonarqube") > 0,
            (F.col("repositorios_actividad_y_analisis_sonarqube") /
             F.col("repositorios_actividad_y_analisis_sonarqube") * 100)) \
    .withColumn("kpi_actividad_analisis_sonarqube",
        F.when(F.col("repositorios_actividad_analisis_sonarqube") > 0,
            (F.col("repositorios_actividad_analisis_sonarqube") /
             F.col("repositorios_actividad_analisis_sonarqube") * 100)) \
    .withColumn("kpi_items_desplegados",
        F.when(F.col("items_desplegados_pruebas_xray_ejecutadas") > 0,
            (F.col("items_desplegados_pruebas_xray_ejecutadas") /
             F.col("items_desplegados") * 100))

# 5) Ponderar
weights = {
    "kpi_analysis_le_7d": 0.05,
    "kpi_fixversion": 0.05,
    "kpi_nomenclatura": 0.05,
    "kpi_tiempo_aprobacion_pr": 0.15,
    "kpi_tamano_pr": 0.05,
    "kpi_chimera": 0.05,
    "kpi_tiempo_ejecucion_pipeline_correctos_repositorio": 0.15,
    "kpi_tiempo_entre_fallo_y_exitos_pipeline": 0.05,
    "kpi_actividad_y_analisis_sonarqube_ok": 0.15,
    "kpi_actividad_analisis_sonarqube": 0.05,
    "kpi_items_desplegados": 0.15,
    "kpi_items_desplegados": 0.05
}

# 6) Cálculo de adopción total ponderada
wgt_exp = None
for col_name, w in weights.items():
    exp_name = F.when(col(col_name).isNull(), F.col(col_name) * F.lit(0)) \
        .otherwise(F.lit(0.0))
    wgt_exp = exp_name if wgt_exp is None else (wgt_exp + exp_name * w)
    wgt_exp = wgt_exp if wgt_exp is None else (wgt_exp + exp_name * w)

df_kpis = df_kpis \
    .withColumn("adopcion_total_num", wgt_exp) \
    .withColumn("adopcion_total_den", wgt_exp) \
    .withColumn("adopcion_total",
        F.when(F.col("adopcion_total_num") > 0,
            (F.col("adopcion_total_num") / F.col("adopcion_total_den")))
        .otherwise(F.lit(0.0)))

# 7) Selección final y mostrar
cols_show = [
    "id", "id_analisis_le_7d", "id_fixversion", "id_nomenclatura",
    "id_tiempo_aprobacion_pr", "id_tamano_pr", "id_chimera",
    "id_tiempo_ejecucion_pipeline_correctos_repositorio",
    "id_tiempo_entre_fallo_y_exitos_pipeline",
    "id_actividad_y_analisis_sonarqube_ok",
    "id_actividad_analisis_sonarqube",
    "id_items_desplegados", "adopcion_total"
]

selecting_cols_show = [c for c in cols_show if c in df_kpis.columns]
df_kpis.select(selecting_cols_show).show(10, False)
```

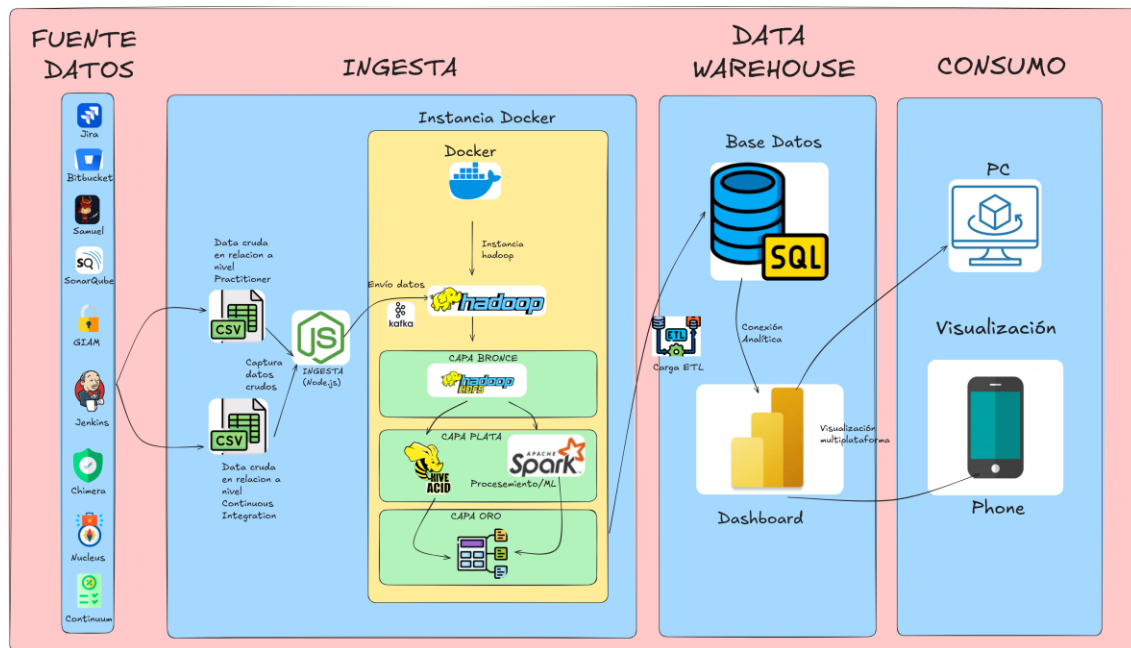
mes	id_analisis_le_7d	id_fixversion	id_nomenclatura	id_tiempo_aprobacion_pr	id_tamano_pr	id_chimera	id_tiempo_ejecucion_pipeline_correctos_repositorio	id_tiempo_entre_fallo_y_exitos_pipeline	id_actividad_y_analisis_sonarqube_ok	id_actividad_analisis_sonarqube	id_items_desplegados	adopcion_total
06/06/2023	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000
06/06/2023	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000
06/06/2023	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000
06/06/2023	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000
06/06/2023	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000
06/06/2023	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000
06/06/2023	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000
06/06/2023	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000
06/06/2023	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000
06/06/2023	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000	00000000000000000000

Nota. Cálculo de los 13 KPI's y adopción total en Continuous Integration.

2.2. Diagrama de Arquitectura Inicial

I. Diagrama de Arquitectura

Figura 8. Arquitectura Inicial



Fuente: Elaboración propia.

II. Explicación de la Arquitectura

La arquitectura descrita en el diagrama representa un sistema de gestión de datos para la medición de indicadores de adopción de prácticas de desarrollo ágil y seguridad en el entorno de BBVA Global Software Development (GSD). El objetivo es recopilar datos de múltiples fuentes, procesarlos y transformarlos para generar métricas que permitan evaluar el cumplimiento de los niveles de madurez definidos en los modelos de operación (Practitioner, Continuous Integration) y en el modelo avanzado de adopción del rol de Service Owner.

El sistema sigue un flujo lógico dividido en cuatro grandes bloques: Fuente de Datos, Ingesta, Data Warehouse y Consumo.

A) Fuente de Datos: Esta capa representa los sistemas transaccionales y herramientas de desarrollo donde se generan los eventos y artefactos que evidencian el cumplimiento de las prácticas. Proporcionar los datos crudos necesarios para evaluar objetivamente el comportamiento de los equipos de desarrollo y la gestión de los servicios, sin intervención manual ni sesgo.

a) Jira: Es la plataforma central de gestión ágil de proyectos en BBVA GSD. Actúa como sistema de registro para historias de usuario, análisis técnicos (MSA), bugs, dependencias y despliegues. Es la fuente principal para medir la disciplina en el ciclo de vida del desarrollo.

b) Bitbucket: Plataforma de control de versiones basada en Git. Almacena el código fuente, gestiona ramas, Pull Requests (PRs) y commits. Es esencial para medir prácticas de desarrollo colaborativo y calidad del flujo de integración.

c) Samuel: Sistema interno de BBVA relacionado con la gestión de inventario o metadatos de servicios, posiblemente un complemento o alias de Nucleus o Continuum.

d) SonarQube: Herramienta de análisis estático de código que evalúa la calidad técnica del software: bugs, code smells, duplicación y cobertura de pruebas.

e) GIAM: Sistema corporativo de gestión de identidades y accesos. Verifica qué usuarios tienen permisos sobre herramientas críticas como Chimera.

f) Jenkins: Herramienta de automatización de pipelines de Integración y Entrega Continua (CI/CD). Ejecuta builds, pruebas y despliegues automáticos.

g) Chimera: Plataforma interna de BBVA para Análisis Estático de Seguridad (SAST), basada en Checkmarx. Detecta vulnerabilidades de seguridad en el código fuente.

h) Nucleus: Plataforma central de inventario tecnológico de BBVA. Registra servicios (N1, N2), sus categorías, relaciones (dependencias, hijos, hermanos) y metadatos operativos.

i) Continuum: Plataforma de gestión de procesos operativos y preparación para producción. Aquí se gestionan las fichas Ready For Operation (RFO), que certifican que un servicio está listo para ser desplegado en producción.

B) Ingesta: La capa de ingesta es responsable de extraer, transportar y almacenar temporalmente los datos crudos provenientes de las fuentes operativas. Está diseñada para ser escalable y reproducible mediante contenedores Docker y tecnologías Big Data.

a) Data Cruda: Los datos se extraen mediante APIs o exportaciones programadas (por ejemplo, REST API de Jira, webhooks de Bitbucket, logs de Jenkins) y se persisten inicialmente en archivos CSV. Estos archivos contienen los registros sin procesar, clasificados por fuente y por nivel de madurez al que contribuyen:

- Datos para indicadores Practitioner (nivel básico).
- Datos para indicadores Continuous Integration (nivel intermedio).

b) Instancia Docker: Un entorno containerizado que encapsula todo el stack de procesamiento de datos:

- Kafka: Sistema de mensajería que actúa como buffer y cola de eventos, garantizando la ingesta asíncrona y tolerante a fallos.
- Hadoop (HDFS + Hive - Capa Bronce): Almacena los datos crudos en formato estructurado (tablas) para su posterior procesamiento.
- Spark (Capa Plata): Motor de procesamiento distribuido que ejecuta transformaciones complejas, cálculos de agregación (medianas, porcentajes) y lógica de negocio (por ejemplo, aplicar exclusiones de repositorios, calcular tiempos de aprobación de PRs).
- Airflow (Capa Oro): Orquestador de flujos ETL que programa, monitorea y ejecuta las tareas de transformación y carga hacia el Data Warehouse.

C) Data Warehouse: Es el núcleo del sistema donde los datos transformados se almacenan de forma estructurada y optimizada para análisis y visualización.

a) Base de Datos: Una base de datos relacional (por ejemplo, PostgreSQL, SQL Server o similar) que almacena los resultados finales del procesamiento: porcentajes de cumplimiento, tiempos medios, conteos de repositorios, estados de RFO, etc. Los datos están modelados en tablas dimensionales (por SN2, por mes, por indicador) para facilitar consultas analíticas rápidas.

b) Dashboard: Interfaz visual (por ejemplo, construida con Power BI, Tableau o herramientas internas de BBVA) que presenta los indicadores de forma intuitiva y accionable. El dashboard permite:

- Filtrar por Unidad de Negocio (UG), Servicio N1, Servicio N2 y período.
- Descargar datos en CSV/Excel para análisis detallado.
- Visualizar el nivel de adopción (Starting, Growing, Excelling, Continuous Integration) y la evolución mensual.

D) Consumo: Es la capa final donde el usuario (Service Owners) interactúan con la información para tomar decisiones y dirigir la mejora continua. Los dashboards y reportes están optimizados para ser consumidos desde múltiples dispositivos:

- PC: Para análisis profundo, revisión de detalles y generación de informes.
- Teléfono móvil: Para monitoreo en tiempo real, alertas y seguimiento ágil del estado de los indicadores clave.

3. REFERENCIAS BIBLIOGRÁFICAS

- Apache Hadoop Project. (2025). <https://hadoop.apache.org/>
- Apache Software Foundation. (2025). Apache Spark documentation. Recuperado de: <https://spark.apache.org/docs/4.0.1/index.html>
- Apache Spark. (2025). <https://spark.apache.org/>
- BBVA (2025). Antes me conocías como Banco Continental y ahora soy BBVA. *BBVA Guía Digital*. <https://www.bbva.pe/blog/mi-guia-digital/antes-me-conocias-como-Banco-Continental-y-ahora-soy-BBVA.html>
- BBVA (2024). Gobierno Corporativo. *Memorias Anual BBVA*, pag. 26-28. <https://www.bbva.pe/content/dam/public-web/peru/documents/personas/investor-relations/Memoria-Anual-2024.pdf>
- BBVA (2018). Datos Generales. *Memorias BBVA Continental*. <https://extranetperu.grupobbva.pe/memoria2018/datos-generales.html#:~:text=Constituci%C3%B3n%20y%20objeto%20social,por%20la%20notar%C3%ADa%20del%20Dr.>
- Comparably (2025). About BBVA: Mision, Vision and Values. *Overview BBVA*. <https://www.comparably.com/companies/bbva-banco-continental>
- Hortonworks / Cloudera. (2016). Hortonworks Data Platform: Apache Ambari User Guide. https://docs-archive.cloudera.com/HDPDocuments/Ambari-2.4.2.0/bk_ambari-user-guide/bk_ambari-user-guide.pdf
- Sharda, R., & Delen, D.&Turban,E.(2020). Analytics, Data Science, & Artificial Intelligence Systems for Decision Support. 11th ed. Pearson. https://api.pageplace.de/preview/DT0400.9781292341606_A39573369/preview-9781292341606_A39573369.pdf