DATASHEET HOUNDOC

Houndoc - Datasheet V5

Noviembre 2026

Contenido

1. Qué es Houndoc 5
2. Arquitectura de la plataforma 5

© Blend 2026. Todos los derechos reservados.

2.1. Integración con fuentes de datos 6
2.2. Estructura de almacenamiento y procesamiento 6
2.3. Capas funcionales 6
2.4. Compatibilidad y despliegue 6
2.5. Seguridad y escalabilidad 7
2.6. Mapa de módulos 7
2.7. Modelos de despliegue y regiones 7
3. Módulos base de la plataforma 8
3.1. IAM - Identity and Access Management 8
Autenticación Federada y Soporte de Protocolos Estándar 8
Gestión de Usuarios y Roles 8
Interoperabilidad con el Ecosistema Houndoc 9
Seguridad y Auditoría 9
3.2. WFM - Data Workflows Management 9
3.3. Data Ingestion Management 10
Arquitectura Técnica y Servicios Asociados 10
3.4. Data Lake 11
Arquitectura Técnica del Componente 11
Beneficios estratégicos del componente 12
3.5. ICG - Information Catalog & Governance 13
Infraestructura Técnica 13
Capacidades del Módulo 13
3.6. AIG - Generative AI 14
Infraestructura Técnica y Amazon Bedrock 15
Ventajas de la infraestructura Bedrock 17
Capacidades Técnicas del Módulo 17
3.7. AAN - Advanced Analytics 19
Procesamiento de Lenguaje Natural y Texto 20
Análisis de Imágenes y Video 21
Modelado Predictivo y Recomendación 21
Entrenamiento y Despliegue de Modelos Propios 22
3.8. Dashboards 22
Infraestructura Técnica y Servicios Asociados 23
Capacidades Técnicas del Módulo 23
4. Capacidades de datos, analítica e IA 24
4.1. SMA - Subscription Management 25

© Blend 2026. Todos los derechos reservados.

Subscription Billing 25
Payment Platform 25
Admin User Interface 25
4.2. MAN - Maps Analytics 26
Visualización Geoespacial Dinámica 26
Integración con Datos Internos y Externos 26
Análisis Espacial y Comparativo 26
Interfaz y Experiencia de Usuario 27
Componentes Técnicos 27
Arquitectura General con AWS 27
4.3. BAO - Business Automation and Optimization 28
Arquitectura del Módulo 29
Integración con la Plataforma 29
4.4. AIS - AI Search Engine 29
Capacidades Técnicas Clave 30
Seguridad y Auditoría 31
Configuración y Despliegue 31
Arquitectura según el entorno de despliegue 31
Funcionamiento de Elasticsearch: fundamentos técnicos 32
Aplicaciones Típicas 32
4.5. WebScrapping 32
Arquitectura de Referencia sobre AWS 33
Componentes Técnicos 34
4.6. AIF - Automated Intelligent Forms 34
Estructura Técnica del Componente 35
Capacidades Adicionales 35
4.7. NSB - Nadia Smart Buddy 36
Arquitectura Técnica y Capacidades 36
Integración con capacidades cognitivas 37
Implementación Multiplataforma 37
Capacidades Técnicas 37
4.8. Scientific Notebooks 39
Arquitectura Técnica 39
Capacidades Extendidas 39
4.9. Servicios complementarios en la nube o entorno local 40
4.9.1. Extracción de Texto y Reconocimiento Óptico de Caracteres (OCR) 40

© Blend 2026. Todos los derechos reservados.

4.9.2. Transcripción Automática de Audio 40
4.9.3. Procesamiento de Lenguaje Natural (NLP) 40
4.9.4. Detección de Entidades y Patrones 41
4.9.5. Pipelines Multimedia Automatizados 41
4.9.6. Migración de Datos y Documentos 41
4.9.7. Personalización e Integración Extendida 42
4.10. Beneficios en datos, analítica e IA 42
Eficiencia Operativa y Automatización Inteligente 43
Aceleración de Proyectos de IA y Analítica 43
Escalabilidad Transparente y Arquitectura Serverless 43
Analítica de Alta Precisión y Modelado Predictivo 43
Seguridad Avanzada y Cumplimiento Normativo 43
Reducción de Errores y Mejora en la Confiabilidad 43
Integración Abierta y Sin Fricciones 44
Flexibilidad Modular y Escenarios Personalizables 44
Alineación Estratégica y Retorno Acelerado 44
4.11. Transformación digital acelerada 44
5. Capacidades de migración y modernización 45
5.1. Houndoc como acelerador de migraciones 45
5.2. El reto de migrar y modernizar 45
5.3. Cómo funciona: agentes de IA con aprobación humana 46
5.4. Capacidades (HM) 46
HM-1. Descubrimiento automatizado 46
HM-2. Diseño de arquitectura como código 47
HM-3. Planeación de olas 48
HM-4. Ejecución y replicación 48
HM-5. Validación automatizada 49
HM-6. Migración de datos y contenido 49
HM-7. Documentación y transferencia 50
HM-8. Modernización de aplicaciones 50
5.5. Fundamentos de aceleración (HF) 51
5.6. Houndoc en cada fase del proyecto de migración 52
5.7. Cargas, plataformas y servicios que cubre 54
5.8. Módulos de migración y modernización 55
5.9. Arquitectura de las capacidades de migración 57
5.10. Modelo de entrega y regiones (HE) 58

© Blend 2026. Todos los derechos reservados.

5.11. Beneficios en migración y modernización 58
Plazos que se cumplen 58
Menos riesgo en cada corte 59
Costo de nube acorde al consumo real 59
Resultados consistentes y auditables 59
Infraestructura como código desde el primer día 59
Un camino de modernización en la misma plataforma 59
Continuidad hacia la explotación de los datos 59
6. Seguridad, privacidad y gobierno 59
7. Propiedad intelectual 60
8. Referencias 60

Houndoc - Plataforma de IA para datos, analítica y migración a la nube

Versión: V5 | **Fecha:** Noviembre 2026 | **Blend**

1. Qué es Houndoc

Houndoc es la plataforma de inteligencia artificial de Blend, construida sobre servicios nativos de Amazon Web Services (AWS). Sobre una base común de módulos de identidad, orquestación, datos, gobierno e inteligencia artificial, Houndoc:

- **Transforma información en conocimiento.** Consolida información dispersa, no estructurada o difícil de procesar, y la explota mediante búsqueda inteligente, analítica avanzada, IA generativa, asistentes conversacionales y automatización de procesos.
- **Acelera la migración y la modernización a la nube.** Automatiza el análisis, el diseño, la planeación, la ejecución, la validación y la documentación de la migración de cargas de trabajo hacia AWS, sin modificar las aplicaciones migradas, y acelera la modernización de aplicaciones y bases de datos.

Houndoc es modular: cada proyecto activa solo los módulos que necesita, y todos comparten la misma base tecnológica y los mismos principios de seguridad. La información que una organización migra a la nube con Houndoc queda catalogada y lista para explotarse con sus capacidades de búsqueda, analítica e IA.

2. Arquitectura de la plataforma

La arquitectura de Houndoc ha sido concebida bajo un enfoque de diseño modular, distribuido y orientado a servicios, lo que garantiza su adaptabilidad a diferentes entornos tecnológicos y escalabilidad en función del volumen de datos, complejidad operativa y necesidades analíticas de cada organización. Esta arquitectura permite su

© Blend 2026. Todos los derechos reservados.

despliegue en escenarios completamente en la nube, en instalaciones locales (on-premises) o bajo esquemas híbridos, según los requerimientos de seguridad, conectividad o normatividad del cliente.

2.1. Integración con fuentes de datos

Uno de los pilares de Houndoc es su capacidad para conectarse con múltiples fuentes de datos estructuradas y no estructuradas, tanto internas como externas. A través de conectores especializados, servicios de ingesta y mecanismos de integración estandarizados, la solución puede interactuar con bases de datos empresariales, sistemas transaccionales, servicios web, repositorios documentales, APIs públicas, plataformas de almacenamiento en la nube, y orígenes no convencionales como sitios web o medios audiovisuales. Este enfoque garantiza una visión completa y consolidada de la información institucional.

2.2. Estructura de almacenamiento y procesamiento

Los datos extraídos son almacenados en un entorno compuesto por un lago de datos y un catálogo de datos. Esta estructura se apoya en formatos optimizados para lectura intensiva, indexación distribuida y consultas de baja latencia.

El procesamiento de la información se realiza a través de flujos orquestados, que permiten la transformación, curaduría, enriquecimiento y análisis automatizado de los datos. La solución es capaz de ejecutar operaciones por lotes (batch processing), procesamiento en tiempo real (streaming) y procesamiento asincrónico de cargas de trabajo intensivas.

2.3. Capas funcionales

La arquitectura se organiza en capas bien definidas que separan responsabilidades y facilitan el mantenimiento y evolución del sistema:

- **Capa de presentación:** Proporciona interfaces gráficas amigables e intuitivas para la interacción de los usuarios con los datos y funcionalidades del sistema.
- **Capa lógica:** Contiene la lógica de negocio, motores de reglas, automatización de procesos, orquestadores de flujo y validadores de integridad.
- **Capa de datos:** Administra el almacenamiento estructurado, no estructurado y semiestructurado, soportando acceso eficiente, versiones históricas y trazabilidad completa.
- **Capa cognitiva:** Alberga modelos de IA y algoritmos de aprendizaje automático para tareas como clasificación, extracción de entidades, análisis de sentimientos, resumen automático, búsqueda inteligente, y generación de contenido.

2.4. Compatibilidad y despliegue

Houndoc ha sido diseñado para operar sobre infraestructuras modernas basadas en contenedores, permitiendo una alta portabilidad, rápida escalabilidad y despliegue flexible en distintos entornos (público, privado o híbrido). La solución se apoya en

© Blend 2026. Todos los derechos reservados.

prácticas de integración y entrega continua (CI/CD), facilitando la evolución incremental y la incorporación segura de nuevas funcionalidades.

2.5. Seguridad y escalabilidad

La plataforma cuenta con mecanismos de seguridad integrados desde el diseño (security by design), incluyendo autenticación federada, control de acceso basado en roles, encriptación de datos en tránsito y en reposo, auditoría completa de acciones y monitoreo en tiempo real de eventos críticos. La arquitectura distribuida permite escalar horizontalmente para gestionar desde cientos hasta millones de registros y documentos sin comprometer el rendimiento.

2.6. Mapa de módulos

Grupo	Módulos	Descripción
Base de la plataforma (sección 3)	IAM, WFM, Data Ingestion Management, Data Lake, ICG, AIG, AAN, Dashboards	Identidad y acceso, orquestación de flujos, ingesta, almacenamiento, gobierno de datos, IA generativa, analítica avanzada y visualización. Son comunes a todas las capacidades
Datos, analítica e IA (sección 4)	SMA, MAN, BAO, AIS, WebScrapping, AIF, NSB, Scientific Notebooks	Suscripciones, análisis geoespacial, automatización de procesos, búsqueda inteligente, extracción web, formularios inteligentes, asistente conversacional y cuadernos científicos
Migración y modernización (sección 5)	MWS, MDA, NLZ, WPL, SMO, MVA, DBM, MOD, MKD	Espacio de trabajo, descubrimiento y evaluación, red y Landing Zone, planeación de olas, migración de servidores, validación, migración de bases de datos, modernización de aplicaciones, y conocimiento y documentación

2.7. Modelos de despliegue y regiones

- Houndoc se despliega en la nube pública, en instalaciones locales (on-premises) o en esquemas híbridos, según los requisitos de seguridad, conectividad y normatividad del cliente.
- En proyectos de migración, Houndoc se ejecuta en la infraestructura de nube de Blend, en las regiones de AWS de **México (Central)** y **Norte de Virginia**, y en otras regiones de AWS que ofrecen los servicios en los que se apoya. Cada

© Blend 2026. Todos los derechos reservados.

proyecto usa la región que defina el cliente según sus requisitos de residencia de datos y normatividad.

3. Módulos base de la plataforma

Los módulos base proveen las capacidades comunes sobre las que se construyen todas las capacidades de Houndoc. Al final de cada módulo se indica su uso específico en migración.

3.1. IAM - Identity and Access Management

El componente **Identity and Access Management (IAM)** en Houndoc proporciona un sistema completo para la autenticación de usuarios, autorización basada en roles y gestión de identidades digitales en entornos empresariales. Diseñado sobre principios de seguridad de identidad federada, este módulo garantiza un acceso seguro, escalable y flexible para usuarios internos, externos y servicios integrados.

IAM está orientado a resolver de manera integral los desafíos asociados al control de acceso en plataformas multiusuario, multicliente y multirol, permitiendo gestionar de manera centralizada las políticas de autenticación y permisos sobre cada uno de los componentes de la plataforma.

Autenticación Federada y Soporte de Protocolos Estándar

IAM soporta autenticación federada mediante protocolos estándares como OpenID Connect, OAuth 2.0 y SAML 2.0 permitiendo la integración con servicios de identidad existentes en la organización (por ejemplo, directorios corporativos, sistemas de gobierno digital o proveedores de identidad en la nube). También permite implementar autenticación mediante usuarios locales, lo que es útil en entornos controlados o en fases de despliegue inicial.

Gestión de Usuarios y Roles

El sistema permite crear, modificar, suspender o eliminar usuarios y asignarles roles definidos con granularidad. Cada rol está asociado a un conjunto de permisos que determinan qué operaciones puede realizar el usuario, sobre qué entidades, y en qué contexto. Esto permite establecer jerarquías de control, políticas de separación de funciones y segmentación lógica por áreas funcionales, organizaciones o unidades operativas.

IAM incluye funcionalidades como:

- Gestión de grupos, dominios y niveles de acceso.
- Integración con directorios corporativos mediante estándares como LDAP.
- Control de sesiones activas y cierre forzado de sesiones en caso de incidentes.
- Registro de eventos relacionados con accesos, autenticaciones fallidas y cambios de privilegios.

© Blend 2026. Todos los derechos reservados.

Interoperabilidad con el Ecosistema Houndoc

IAM se integra de forma nativa con todos los componentes de la plataforma, habilitando o restringiendo funcionalidades según el perfil del usuario autenticado. También puede funcionar como proveedor de identidad para aplicaciones o módulos satélites, consolidando la experiencia de autenticación en un único punto de entrada (single sign-on).

Seguridad y Auditoría

Todo el acceso gestionado por IAM queda registrado en una bitácora de auditoría que almacena información crítica como usuario, acción, componente afectado, fecha, dirección IP y resultado de la operación. Esta información es clave para garantizar la trazabilidad, la detección temprana de comportamientos anómalos y el cumplimiento de normativas de seguridad y protección de datos.

Uso en migración. Define quién puede consultar, aprobar o ejecutar cada fase de la migración, y deja evidencia de cada aprobación.

3.2. WFM - Data Workflows Management

El componente **WFM** es el orquestador técnico de flujos de datos dentro de la plataforma Houndoc. Su función principal es coordinar, ejecutar y supervisar procesos de transformación, carga y preparación de datos de manera automatizada y desacoplada. Está construido sobre una arquitectura serverless que garantiza escalabilidad, alta disponibilidad y ejecución resiliente de flujos en entornos distribuidos.

WFM combina varios servicios nativos de AWS especializados en la gestión de pipelines de datos:

- **AWS Glue** es utilizado como motor principal de transformación. Permite ejecutar trabajos ETL (extracción, transformación y carga) a partir de fuentes estructuradas o semiestructuradas. Su integración con el catálogo de datos permite la inferencia de esquemas y el procesamiento de datos almacenados en Amazon S3 de forma automática o programada.
- **AWS Step Functions** proporciona la capacidad de definir flujos complejos con lógica de control explícita (secuencias, bifurcaciones, condiciones, esperas, paralelismos, reintentos y manejo de errores). Cada paso del flujo puede representar una función Lambda, un job de Glue o una invocación externa, todo orquestado en un grafo dirigido.
- **AWS Lambda** actúa como ejecutor de tareas ligeras y desacopladas dentro del flujo, permitiendo encapsular transformaciones específicas, validaciones, conversión de formatos o disparo de eventos subsecuentes. Su modelo sin servidor habilita la ejecución paralela y bajo demanda, sin provisión de infraestructura.
- **Amazon EventBridge** permite la orquestación reactiva de flujos. Al recibir eventos desde servicios internos o externos, desencadena automáticamente la ejecución de workflows definidos, generando una arquitectura basada en eventos, altamente desacoplada y extensible.

© Blend 2026. Todos los derechos reservados.

- **Amazon CloudWatch** registra métricas detalladas, logs de ejecución y alarmas asociadas a cada uno de los elementos del flujo. Permite observar el rendimiento de tareas específicas, establecer umbrales críticos y auditar en tiempo real el comportamiento del sistema de orquestación.

WFM permite definir los workflows como definiciones declarativas, versionables y auditables. Cada componente del flujo está desacoplado, ejecutado de manera aislada y monitoreado por separado, lo cual garantiza estabilidad operacional y facilita el mantenimiento técnico. La naturaleza serverless de su infraestructura permite adaptar automáticamente la capacidad de procesamiento al volumen de datos y a la complejidad de las operaciones sin intervención manual.

Uso en migración. Gobierna las fases de la migración, los puntos de aprobación humana y la ejecución simultánea de varias olas. Cada flujo de migración es una especificación versionable y auditable.

3.3. Data Ingestion Management

El componente **Data Ingestion Management** de Houndoc permite la adquisición automatizada de datos desde múltiples fuentes heterogéneas, con soporte para esquemas de ingestión por lotes, flujos en tiempo real y transferencia masiva desde entornos locales o externos. Su arquitectura está orientada a garantizar una captación eficiente, validada y adaptable de datos estructurados y no estructurados, provenientes de fuentes internas, terceros o sistemas abiertos.

Arquitectura Técnica y Servicios Asociados

Este componente está construido sobre una infraestructura modular que aprovecha diversos servicios de AWS para gestionar escenarios de ingestión de diversa naturaleza:

- **Carga desde bases de datos estructuradas**

Para replicación desde motores relacionales, se utiliza **AWS Database Migration Service (DMS)**. Este servicio permite ejecutar cargas iniciales o replicación continua desde bases de datos locales o en la nube, asegurando la sincronización de datos sin afectar la disponibilidad del origen.

- **Transferencia masiva de archivos**

Para la ingestión de grandes volúmenes de archivos desde entornos on-premises o redes compartidas, se emplea **AWS DataSync**, que permite copiar datos de forma eficiente hacia Amazon S3, con soporte para control de versiones, programación y validación automática de integridad.

- **Flujos en tiempo real**

En contextos donde los datos deben ser capturados y procesados a medida que ocurren, se implementan arquitecturas basadas en **Amazon Kinesis Data Streams** o **Kinesis Data Firehose**, que permiten capturar, bufferizar y enviar flujos de datos a sistemas de almacenamiento o procesamiento sin pérdida. Este enfoque es ideal para ingestión desde sensores, logs, APIs de eventos o sistemas de monitoreo.

- **Ejecución y preprocesamiento sin servidor**

© Blend 2026. Todos los derechos reservados.

Para tareas ligeras de transformación o validación al momento de la ingestión, se utiliza **AWS Lambda**, lo cual permite ejecutar funciones en respuesta a eventos como la llegada de nuevos archivos, actualización de bases de datos o señales de sistemas externos.

- **Organización en el data lake**

Los datos capturados se almacenan de forma estructurada en Amazon S3, aplicando particionamiento automático por carpetas jerárquicas (por ejemplo, fecha, dominio, tipo de dato), lo cual facilita su descubrimiento, consulta y procesamiento posterior. Esta organización puede ser gestionada y catalogada utilizando definiciones de esquema asociadas al Glue Data Catalog.

El componente permite configurar estrategias de ingestión según la criticidad de los datos, el volumen esperado y la complejidad de integración, asegurando una base confiable, segura y escalable para el procesamiento posterior de la información dentro de la plataforma.

Uso en migración. Incorpora los inventarios, métricas y configuraciones de origen, y ejecuta la transferencia masiva de archivos y la replicación de bases de datos. Para volúmenes que exceden la capacidad de los enlaces se integra con dispositivos de transferencia física de AWS Snow Family.

3.4. Data Lake

El componente **Data Lake** de Houndoc constituye una plataforma robusta diseñada para gestionar eficientemente grandes volúmenes de información, proporcionando una **capa centralizada, persistente y escalable** para la administración de datos en todos sus estados de procesamiento (estructurados, semiestructurados y no estructurados).

Construido bajo una arquitectura desacoplada y alineado con las mejores prácticas tecnológicas, el Data Lake de Houndoc aprovecha la potencia y seguridad de los servicios nativos de **Amazon Web Services (AWS)**, asegurando una operación eficiente, duradera, segura y altamente disponible.

Arquitectura Técnica del Componente

Almacenamiento basado en Amazon S3 El núcleo del Data Lake utiliza **Amazon S3** como infraestructura primaria, organizado en un esquema jerárquico compuesto por tres zonas lógicas claramente definidas que permiten la segmentación efectiva de los datos según su madurez operativa:

- **Raw**: Almacenamiento inicial de datos en estado original, tal como se reciben desde la fuente de origen.
- **Cleaned**: Almacenamiento intermedio con datos normalizados, validados y preparados para procesamiento posterior.
- **Curated**: Almacenamiento final con datos enriquecidos, transformados, estructurados y listos para consumo analítico o aplicación de modelos de inteligencia artificial.

© Blend 2026. Todos los derechos reservados.

Cada zona está estructurada mediante un sistema avanzado de **particionamiento jerárquico** basado en atributos, optimizando así las consultas, reduciendo tiempos de escaneo y minimizando costos operativos.

Formatos de almacenamiento optimizados El almacenamiento utiliza formatos eficientes como **Parquet** y **ORC** para datos estructurados y semiestructurados, que permiten:

- **Acceso eficiente únicamente a las columnas requeridas en cada consulta**, acelerando la recuperación de datos y optimizando el rendimiento.
- **Facilidad de integración y alto rendimiento en herramientas de análisis masivo y procesamiento distribuido**, simplificando su uso en diferentes plataformas analíticas.

Adicionalmente, se emplea almacenamiento estructurado en **JSON**, asegurando compatibilidad con motores de búsqueda y aplicaciones de búsqueda full-text.

Procesamiento distribuido y escalable Las tareas de **Extracción, Transformación y Carga (ETL)** se gestionan mediante **AWS Glue**, aprovechando procesos automatizados escritos en lenguaje **PySpark**. Glue opera sobre infraestructura serverless, realizando tareas complejas tales como limpieza exhaustiva, validación estructural, deduplicación y enriquecimiento semántico.

Orquestación inteligente de flujos con AWS Step Functions Para asegurar precisión y coherencia operativa, los procesos de ingestión, transformación y publicación se coordinan a través de **AWS Step Functions**, permitiendo controlar detalladamente cada paso del flujo de procesamiento mediante:

- Ejecución condicional lógica.
- Gestión automatizada de errores.
- Coordinación secuencial y paralela de actividades críticas.

Expansión hacia arquitectura Lakehouse Este componente también está preparado para expandirse hacia un modelo de arquitectura tipo **Lakehouse**, mediante la incorporación de herramientas analíticas adicionales como **Amazon Redshift Spectrum**, que permite realizar consultas analíticas SQL directamente sobre datos almacenados en **Amazon S3**, o bien a través de instancias dedicadas de **Amazon Redshift**, alcanzando así una integración efectiva entre la flexibilidad de almacenamiento del Data Lake y la alta capacidad analítica característica de los almacenes de datos columnar. Esto habilita análisis avanzados y de alto rendimiento, especialmente útiles para cargas analíticas complejas y dinámicas.

Beneficios estratégicos del componente

- **Escalabilidad y flexibilidad**: Capacidad para manejar volúmenes crecientes de datos diversos de manera eficiente y ajustable a diferentes necesidades de negocio.
- **Optimización operativa**: Reducción de costos de almacenamiento y mejora significativa en tiempos de respuesta de consultas analíticas.
- **Alta integración con AWS**: Aprovechamiento total del ecosistema AWS para procesos eficientes y simplificados, garantizando seguridad y disponibilidad.

© Blend 2026. Todos los derechos reservados.

Uso en migración. Guarda el modelo de conocimiento del proyecto (inventario, dependencias, métricas, diseños, olas, decisiones y evidencia) y, en migraciones de plataformas analíticas, sirve como destino de los datos migrados.

3.5. ICG - Information Catalog & Governance

El componente **Information Catalog & Governance (ICG)** de Houndoc habilita la catalogación, trazabilidad, control de acceso y gobernanza de los activos de datos administrados por la plataforma. Su propósito es proporcionar visibilidad estructurada sobre el universo de información disponible, establecer políticas de control y facilitar el cumplimiento de normas de calidad, privacidad y uso responsable de los datos.

ICG está diseñado para operar como un catálogo de datos corporativo, centralizando el descubrimiento, clasificación, linaje y control de acceso sobre datasets estructurados y no estructurados. Funciona como capa transversal sobre el data lake y otros componentes de almacenamiento, conectando la infraestructura de datos con las reglas de negocio y los principios de gobierno institucional.

Infraestructura Técnica

El módulo se basa en servicios nativos de AWS orientados al gobierno de datos, destacando:

- **AWS Glue Data Catalog**

Actúa como el catálogo central de metadatos de Houndoc. Cada tabla, dataset o colección de datos registrada en el data lake es representada como una entrada en el catálogo, que incluye información técnica (esquema, ubicación, formato), lógica (dominio, propietario, frecuencia de actualización) y semántica (tipo de dato, confidencialidad, uso permitido). Esto permite descubrir y consultar metadatos de forma programática y desde interfaces gráficas.

- **AWS Lake Formation**

Administra las políticas de acceso y gobernanza sobre los datos almacenados en Amazon S3. Permite definir permisos finos sobre columnas, filas y tablas completas, en función del rol del usuario, el dominio de datos o la sensibilidad de la información. Además, soporta el registro de transacciones, el control de versiones de los datos y el linaje desde origen hasta consumo.

Capacidades del Módulo

- **Descubrimiento de Datos**

Permite a los usuarios explorar los datos disponibles mediante búsqueda por nombre, categoría, etiquetas, dominio funcional o tipo de dato. Las entradas del catálogo pueden enriquecerse con descripciones técnicas, funcionales y enlaces a ejemplos de uso.

- **Clasificación Automática y Etiquetado**

ICG incluye funciones de detección automática de información sensible (como datos personales o financieros) y sugiere clasificaciones de acuerdo con políticas predefinidas. Los administradores pueden aplicar etiquetas personalizadas para facilitar la segmentación y auditoría.

© Blend 2026. Todos los derechos reservados.

- **Linaje y Trazabilidad**

Registra el flujo completo de los datos desde su origen (sistemas fuente, APIs, cargas manuales), pasando por procesos de transformación, hasta su visualización o exportación. Este linaje es esencial para auditar el uso de los datos, comprender impactos de cambios o responder a requerimientos regulatorios.

- **Gestión de Políticas de Acceso**

El módulo permite definir políticas de acceso basadas en roles, niveles de clasificación y condiciones contextuales (como origen de la solicitud, horario o dominio). Estas políticas son evaluadas dinámicamente y aplicadas a través de los servicios de seguridad y gobernanza integrados.

- **Auditoría y Cumplimiento**

Toda interacción con el catálogo (consultas, modificaciones, asignación de permisos) queda registrada en logs estructurados, que pueden ser analizados para verificar cumplimiento normativo, identificar usos indebidos o auditar la evolución de los activos de datos.

ICG fortalece la arquitectura de datos de Houndoc al permitir que los activos informacionales no solo estén disponibles, sino también gestionados de forma ordenada, segura y alineada con los principios de gobernanza institucional.

Uso en migración. Clasifica la información sensible antes de migrarla y registra el linaje de lo migrado para auditorías y requisitos de protección de datos.

3.6. AIG - Generative AI

El componente **Generative AI (AIG)** de Houndoc integra modelos fundacionales de **inteligencia artificial generativa** de última generación, capaces de procesar una variedad de tipos de datos, tales como texto, imágenes, audio y datos estructurados. AIG permite la creación automática y contextual de contenido, generando no solo texto, sino también imágenes y videos.

AIG utiliza capacidades avanzadas de inteligencia artificial generativa para llevar a cabo diversas tareas cognitivas. Estas incluyen el análisis de sentimiento, que permite evaluar emociones o intenciones en textos o audios, lo cual resulta especialmente útil para el análisis de feedback o la atención al cliente. También se puede realizar extracción de información, identificando y extrayendo entidades clave de grandes volúmenes de documentos, lo que facilita la organización y análisis de datos no estructurados. AIG también tiene la capacidad de generar respuestas automáticas a preguntas basadas en los datos disponibles, lo que agiliza la interacción con los usuarios o la automatización de procesos. Además, AIG es capaz de generar contenido visual, como la creación de subtítulos para imágenes o la generación de gráficos complementarios a partir de descripciones textuales. También realiza reconocimiento de objetos en imágenes y videos, lo que permite la identificación automática de elementos visuales relevantes. Por último, la capacidad de síntesis de texto permite la creación de resúmenes de contenido o la redacción automática de informes completos.

Como resultado de estas capacidades, AIG es capaz de generar diferentes tipos de outputs. Estos incluyen texto, como informes automatizados, resúmenes de contenido

© Blend 2026. Todos los derechos reservados.

largo o respuestas personalizadas a preguntas específicas. También puede generar imágenes, como la creación de imágenes a partir de descripciones textuales o la generación de gráficos complementarios para informes o presentaciones. Además, AIG tiene la capacidad de generar videos, produciendo clips de video basados en descripciones textuales o resúmenes visuales de información compleja.

Este módulo permite a las organizaciones transformar grandes volúmenes de datos no estructurados en contenido útil y procesable, mejorando la eficiencia operativa, la toma de decisiones y la personalización de los servicios. AIG automatiza la creación de contenido y facilita la interacción con los datos de manera dinámica, adaptándose a las necesidades específicas de cada proceso.

AIG está diseñado para ser utilizado en escenarios donde se requiere redacción asistida, generación de lenguaje natural y explicación detallada de resultados. Su capacidad para realizar normalización semántica asegura que los datos se presenten de manera coherente y accesible, facilitando la comprensión incluso en contextos complejos. Además, AIG permite la producción automatizada de documentos, lo que optimiza la creación de informes, contratos, o cualquier otro tipo de contenido extenso, todo ello adaptado a las necesidades del usuario. También es ideal para la generación de recomendaciones personalizadas y la respuesta automatizada a consultas complejas, proporcionando una solución eficaz para la automatización de la toma de decisiones y la interacción con los usuarios, sin perder precisión ni contexto. En resumen, AIG simplifica y agiliza los procesos donde el análisis de datos, la creación de contenido y la generación de insights son fundamentales para la operación eficiente y escalable de las organizaciones.

Infraestructura Técnica y Amazon Bedrock

Este módulo opera sobre Amazon Bedrock, el servicio de AWS que proporciona acceso a modelos fundacionales (Foundation Models, FMs) de varios proveedores líderes en inteligencia artificial, todo a través de una única API. Bedrock elimina la necesidad de aprovisionar o gestionar infraestructura compleja de machine learning, permitiendo que las organizaciones se centren en aprovechar el poder de los modelos de última generación sin tener que preocuparse por los aspectos técnicos de su implementación y escalabilidad. Al estar basado en la infraestructura de AWS, Amazon Bedrock asegura una alta disponibilidad, seguridad y capacidad de escalado, lo que permite a los usuarios obtener resultados precisos y rápidos de forma eficiente y económica.

A través de Amazon Bedrock, AIG puede acceder a una variedad de modelos generativos y de aprendizaje profundo de alto rendimiento, adaptados a diferentes casos de uso. Estos modelos, proporcionados por algunos de los principales actores en la industria de la inteligencia artificial, están optimizados para tareas de procesamiento de lenguaje natural, análisis de imagen, generación de contenido y mucho más, lo que expande las capacidades de AIG a un espectro aún más amplio de aplicaciones.

Modelos compatibles disponibles en AIG vía Bedrock:

- **Amazon Nova:** Amazon Nova está disponible en varias versiones: Micro, Lite y Pro, y acepta entradas de texto, imágenes y videos para generar salidas de texto. Además, las herramientas Nova Canvas y Nova Reels permiten la creación de

© Blend 2026. Todos los derechos reservados.

contenido visual y multimedia, generando imágenes y videos a partir de descripciones textuales, lo que facilita la creación de materiales visuales creativos. Nova Sonic, por su parte, está diseñado para tomar audio como entrada y generar audio y texto como salida, brindando soporte para conversaciones de voz en tiempo real con una rica contextualización, lo que es ideal para aplicaciones de asistentes virtuales o centros de llamadas.

- **Anthropic Claude (Claude 1, 2, 3):** Los modelos Claude están especialmente diseñados para comprender y procesar instrucciones complejas. Estos modelos sobresalen en la generación de lenguaje formal y en el manejo de contextos extensos, lo que los hace ideales para aplicaciones que requieren interacciones largas y detalladas. Los modelos Claude son altamente competentes en tareas como la generación de respuestas coherentes, la interpretación de preguntas complejas y la elaboración de contenidos bien estructurados en diferentes contextos, garantizando un alto nivel de precisión y calidad en la comunicación.
- **AI21 Labs Jurassic:** Estos modelos se especializan en generación creativa y en proporcionar respuestas largas con un lenguaje enriquecido, lo que los hace ideales para generar contenido detallado y narrativas extensas. Este modelo se utiliza en aplicaciones donde se requiere un estilo de lenguaje sofisticado y bien estructurado, como en la redacción de artículos, blogs o contenido creativo, proporcionando respuestas de alta calidad que mantienen la coherencia a lo largo de interacciones largas.
- **Meta Llama 2:** Está diseñado para un alto rendimiento general con un enfoque en bajo costo computacional y un amplio soporte multilingüe. Este modelo está optimizado para tareas como la traducción de idiomas, la interpretación de texto y el análisis de grandes volúmenes de información en múltiples idiomas, lo que lo convierte en una herramienta valiosa para aplicaciones internacionales o multilingües. Además, su eficiencia en el uso de recursos computacionales lo hace adecuado para una amplia gama de tareas en entornos de alta demanda.
- **Stability AI (Stable Diffusion):** Es un modelo de generación de imágenes a partir de texto (text-to-image), desarrollado por Stability AI. Este modelo es capaz de crear imágenes visualmente impresionantes basadas en descripciones textuales, permitiendo a los usuarios generar ilustraciones, gráficos, arte digital y otros contenidos visuales a partir de simples comandos de texto. Esta capacidad es útil en áreas como la publicidad, el diseño gráfico, la creación de contenido multimedia y la personalización de experiencias visuales.
- **DeepSeek-R1:** Es un modelo especializado en la búsqueda y recuperación de información de gran escala. Su principal objetivo es mejorar la eficiencia en la búsqueda de datos dentro de bases de conocimiento extensas, utilizando técnicas avanzadas de procesamiento de lenguaje natural (NLP). Este modelo es ideal para entornos empresariales que requieren consultar grandes cantidades de datos no estructurados, como documentos, correos electrónicos o contenido multimedia, y obtener resultados rápidos y precisos.
- **Amazon Titan**: Son modelos entrenados y optimizados específicamente por AWS para tareas empresariales comunes, como clasificación, resumen de textos y embeddings vectoriales. Estos modelos son ideales para aplicaciones que requieren procesamiento de grandes volúmenes de datos, como la automatización de la gestión de documentos, la clasificación de contenido y la creación de representaciones vectoriales de datos. Titan se destaca por su

© Blend 2026. Todos los derechos reservados.

eficiencia y escabilidad, lo que lo convierte en una opción popular para empresas que buscan integrar inteligencia artificial en flujos de trabajo empresariales.

- **Entre otros modelos.**

Ventajas de la infraestructura Bedrock

- **Escalabilidad sin infraestructura dedicada**

Una de las principales ventajas de Amazon Bedrock es su escalabilidad automática, lo que permite a las organizaciones ejecutar inferencias con una capacidad de respuesta que se adapta a las cargas de trabajo variables, sin necesidad de aprovisionar ni administrar infraestructura dedicada. Esto elimina la complejidad de gestionar clústeres o instancias de cómputo, ya que Bedrock ajusta automáticamente los recursos en función de la demanda. Este enfoque permite a las empresas operar con eficiencia y flexibilidad, asegurando que las aplicaciones impulsadas por IA puedan manejar picos de tráfico o procesos intensivos de forma fluida y sin interrupciones, optimizando tanto los costos como el rendimiento.

- **Flexibilidad en selección de modelos**

Amazon Bedrock ofrece una gran flexibilidad al permitir la selección dinámica de modelos según el caso de uso específico, el tipo de entrada que se maneje (texto, imágenes, audio, etc.), el idioma requerido y la complejidad del resultado esperado. Esto significa que las organizaciones no tienen que comprometerse a un único modelo o proveedor, sino que pueden optar por el modelo más adecuado para cada tarea o aplicación. Además, esta flexibilidad se extiende a la integración, ya que los usuarios no necesitan hacer modificaciones significativas en sus aplicaciones o sistemas al cambiar entre modelos. Esto se traduce en una implementación más ágil y una mejor adaptación a las necesidades específicas de cada proyecto o cliente, sin los típicos problemas de compatibilidad entre modelos o cambios tecnológicos complejos.

- **Seguridad y privacidad empresarial**

Una ventaja crucial de Amazon Bedrock es su robusto enfoque en seguridad y privacidad, particularmente en sectores regulados. Bedrock garantiza que los datos procesados no se utilicen para entrenar modelos adicionales, protegiendo así la información confidencial y asegurando el cumplimiento de normativas sobre privacidad de datos. Además, la infraestructura de Bedrock asegura el aislamiento completo de los datos, lo que significa que los datos procesados no se filtran ni se comparten con otros procesos o clientes. Esto hace que Amazon Bedrock sea una opción segura y viable para industrias como la salud, finanzas y gobierno, donde la privacidad y la protección de datos son requisitos críticos. Esta garantía de privacidad es fundamental para implementar soluciones de IA generativa en entornos altamente regulados, donde la gestión de datos sensibles es una prioridad.

Capacidades Técnicas del Módulo

- **Entrada multimodal**: El módulo AIG de Houndoc soporta una entrada multimodal, permitiendo la recepción de texto, metadatos, listas, tablas o combinaciones estructuradas como insumos para su procesamiento. Esta flexibilidad de entrada facilita la personalización semántica del contenido

© Blend 2026. Todos los derechos reservados.

generado, adaptando la salida según los datos proporcionados y garantizando que los resultados sean precisos y contextualizados. Al aceptar una amplia gama de formatos de entrada, AIG puede integrarse de manera eficiente en diversos sistemas y flujos de trabajo, permitiendo una mayor adaptabilidad a diferentes tipos de datos y necesidades empresariales.

- **Prompt Engineering avanzado**: AIG utiliza técnicas avanzadas de prompt engineering, lo que permite la creación de estructuras de prompt configurables y parametrizadas. Estas estructuras son ajustables según el tipo de usuario, el rol funcional o la lógica de proceso específica del sistema. Esta capacidad otorga un alto nivel de personalización, lo que facilita la adaptación del modelo a diferentes contextos operativos o interacciones. Esto es particularmente útil para organizaciones que necesitan ajustar el comportamiento de la IA según las necesidades particulares de cada departamento o caso de uso, garantizando que los resultados generados estén alineados con los objetivos y el lenguaje específicos de la organización.
- **Personalización con técnicas como fine-tuning o few-shot learning**: AIG permite la personalización avanzada mediante técnicas como fine-tuning y few-shot learning, lo que posibilita la adaptación de los modelos a terminología específica, marcos normativos o estilos lingüísticos propios de la organización. Estas técnicas aseguran que el modelo no solo se ajuste a los requisitos técnicos, sino también a las peculiaridades del dominio en el que se emplea, como términos legales, financieros o médicos. Esta personalización también facilita la integración con normativas locales o internacionales, asegurando que las salidas del modelo sean coherentes con las regulaciones y prácticas propias de la industria.
- **Postprocesamiento de resultados**: El módulo AIG incluye capacidades avanzadas de postprocesamiento, que permiten aplicar validaciones automáticas sobre los resultados generados. Esto incluye la verificación de la estructura, el lenguaje utilizado, los formatos esperados, la longitud del contenido, la presencia de lenguaje inapropiado y la consistencia temática. Este proceso asegura que el contenido final no solo sea relevante y preciso, sino también adecuado para su uso en entornos profesionales y regulatorios, reduciendo errores y mejorando la calidad general de las respuestas generadas.
- **Ejecución embebida en flujos**: AIG puede ser fácilmente embebido en flujos de trabajo automatizados y sistemas de orquestación, lo que permite la generación automática de contenidos dentro de procesos más amplios. Esto incluye tareas como la creación de diagnósticos preliminares, la generación de respuestas normativas, la elaboración de resúmenes de documentos y la construcción de argumentaciones legales. Al estar integrado de manera fluida en flujos de trabajo preexistentes, AIG optimiza la productividad y eficiencia al automatizar tareas que tradicionalmente son manuales y repetitivas, permitiendo a las organizaciones enfocarse en actividades de mayor valor estratégico.

AIG permite incorporar IA generativa de manera segura, escalable y productiva, lo que facilita la transformación de procesos tradicionalmente manuales en sistemas inteligentes, adaptativos y altamente automatizados. Al aprovechar la potencia de modelos avanzados de IA, las organizaciones pueden optimizar sus flujos de trabajo, reducir tiempos de procesamiento y mejorar la precisión y coherencia en la entrega

© Blend 2026. Todos los derechos reservados.

de resultados. Esta capacidad no solo permite una mayor eficiencia operativa, sino que también habilita nuevas formas de interacción con los datos, mejorando la toma de decisiones y permitiendo una personalización más precisa de servicios y productos. Con la integración de AIG, las organizaciones pueden adaptarse rápidamente a cambios del mercado o requisitos regulatorios, transformando desafíos complejos en oportunidades de innovación continua y manteniendo una ventaja competitiva sostenible en un entorno empresarial en constante evolución.

Uso en migración. Es el motor de los agentes de IA: genera los análisis de dependencias y preparación, los diseños y el código de infraestructura, los planes de olas y contingencia, el código modernizado, la documentación y las respuestas del asistente.

3.7. AAN - Advanced Analytics

El componente **Advanced Analytics (AAN)** de Houndoc es el módulo especializado en analítica avanzada de la plataforma, diseñado para integrar diversas técnicas de inteligencia artificial que permiten procesar automáticamente datos estructurados y no estructurados en distintos formatos: texto, documentos, registros tabulares, archivos de audio, imágenes y video.

AAN combina tecnologías como procesamiento de lenguaje natural (NLP), análisis de imagen y video, modelos de clasificación y predicción, sistemas de recomendación y extracción automatizada de entidades clave, proporcionando un conjunto completo de capacidades analíticas al servicio de procesos de auditoría, fiscalización, monitoreo, supervisión o evaluación estratégica.

El módulo facilita la analítica descriptiva, permitiendo a las organizaciones explorar datos históricos, identificar patrones y relaciones significativas, y generar visualizaciones y reportes que mejoran la comprensión de los procesos observados. Esta capa de análisis apoya el diagnóstico situacional, la detección de ineficiencias y la elaboración de tableros de control dinámicos.

Además, AAN habilita analítica predictiva y prescriptiva, utilizando modelos de inteligencia artificial tradicionales que permiten anticipar eventos, evaluar escenarios futuros y recomendar acciones basadas en resultados probabilísticos. Esto refuerza la toma de decisiones con fundamentos analíticos, contribuyendo a una gestión más proactiva, eficiente y orientada a resultados.

Las funcionalidades del módulo se ejecutan sobre una infraestructura cloud segura y escalable, cumpliendo con altos estándares de gobernanza, privacidad y disponibilidad. AAN se integra de manera fluida con los demás componentes de Houndoc, asegurando una experiencia coherente y unificada en la gestión avanzada de datos.

En conjunto, Advanced Analytics transforma los datos en conocimiento accionable, permitiendo a las organizaciones anticiparse a los desafíos, reducir riesgos y optimizar el rendimiento operativo a través del uso estratégico de inteligencia artificial.

El módulo de Analítica Avanzada (AAN) de Houndoc se integra de manera directa con la suite de servicios de inteligencia artificial de AWS, aprovechando cada uno de ellos para abordar tareas específicas de análisis cognitivo y procesamiento automático de

© Blend 2026. Todos los derechos reservados.

información. Estos servicios, impulsados por tecnologías avanzadas de machine learning, permiten optimizar diversos procesos, desde la interpretación de textos hasta la creación de modelos predictivos y la automatización de interacciones. A continuación, se detallan los componentes funcionales principales que conforman esta integración, destacando cómo cada uno contribuye a la mejora de la toma de decisiones y la eficiencia operativa.

Procesamiento de Lenguaje Natural y Texto

- **Amazon Comprehend**: Este componente, integrado dentro del módulo de Analítica Avanzada de Houndoc, permite analizar y comprender el lenguaje natural contenido en documentos o flujos de texto. Mediante técnicas avanzadas de procesamiento de lenguaje natural, identifica automáticamente entidades, temas, sentimientos, relaciones y frases clave, facilitando la estructuración de información no estructurada. También habilita la personalización de modelos para reconocer categorías específicas, adaptándose a contextos sectoriales como legal, financiero o de atención al cliente. Gracias a estas capacidades, mejora significativamente la eficiencia en la extracción de insights y en la toma de decisiones basada en texto.
- **Amazon Textract**: Dentro del módulo de Analítica Avanzada de Houndoc, este componente permite extraer automáticamente texto y datos estructurados a partir de documentos escaneados, incluyendo formularios y tablas. No se limita a reconocer caracteres, sino que comprende la disposición lógica del contenido, habilitando la digitalización eficiente de documentos físicos y la automatización de flujos operativos. Su precisión en la extracción de información crítica lo hace especialmente útil en sectores que requieren procesar grandes volúmenes de documentación, como salud, legal o financiero. Esta funcionalidad basada en inteligencia artificial optimiza tiempos de procesamiento y contribuye a decisiones más ágiles y fundamentadas.
- **Amazon Polly**: Este componente forma parte del módulo de Analítica Avanzada de Houndoc y se encarga de transformar texto en voz natural con gran realismo, empleando tecnologías de síntesis de voz multilingüe y configurable. Su uso es clave para generar experiencias auditivas personalizadas en aplicaciones que requieren interacción por voz, como asistentes virtuales, sistemas de accesibilidad, centros de contacto o plataformas educativas. Permite ajustar aspectos como el tono, la velocidad y la entonación, adaptándose a diferentes públicos y contextos. Esta funcionalidad, basada en inteligencia artificial, no solo mejora la interacción hombre-máquina, sino que amplía las posibilidades de automatización en la entrega de contenidos y servicios.
- **Amazon Transcribe**: Como parte del módulo de Analítica Avanzada de Houndoc, este componente permite convertir automáticamente audio en texto escrito utilizando modelos de reconocimiento automático del habla basados en inteligencia artificial. Es capaz de procesar grabaciones o flujos de audio en tiempo real, identificando palabras con alta precisión y distinguiendo entre diferentes hablantes cuando es necesario.

Esta funcionalidad resulta especialmente útil en escenarios donde se requiere documentar conversaciones, transcribir reuniones, generar actas, o facilitar el análisis de interacciones en centros de atención, servicios legales o entornos médicos.

© Blend 2026. Todos los derechos reservados.

Además, admite múltiples idiomas y puede integrarse fácilmente en flujos de trabajo automatizados, contribuyendo a optimizar procesos operativos, reducir tiempos de documentación y fortalecer la toma de decisiones basada en evidencias verbales.

- **Amazon Lex**: Este componente, integrado en el módulo de Analítica Avanzada de Houndoc, permite desarrollar agentes cognitivos capaces de interpretar y responder en lenguaje natural, gestionando diálogos complejos y ejecutando acciones automatizadas según la intención del usuario. Utiliza técnicas avanzadas de procesamiento del lenguaje y modelado de diálogo para comprender con precisión lo que se expresa, facilitando la interacción hombre-máquina de manera fluida y contextual.

Cuando se combina con otros componentes del ecosistema, este agente puede desplegarse en múltiples canales y formatos, desde un chat integrado en sistemas web hasta representaciones más sofisticadas como avatares realistas que interactúan por voz. Esta versatilidad permite su aplicación en procesos de atención automatizada, soporte, recopilación de datos o capacitación virtual, aportando eficiencia operativa, reducción de tiempos de respuesta y una experiencia de usuario más enriquecida. La base en inteligencia artificial permite adaptar el comportamiento del agente al contexto específico de cada organización, mejorando significativamente la toma de decisiones y la automatización de tareas repetitivas.

Análisis de Imágenes y Video

- **Amazon Rekognition**: Este componente, como parte del módulo de Analítica Avanzada de Houndoc, permite analizar imágenes y videos para identificar automáticamente elementos, patrones y atributos relevantes mediante técnicas avanzadas de inteligencia artificial. Entre sus capacidades se encuentran el reconocimiento de objetos, escenas, actividades, rostros y expresiones, así como la detección de contenido inapropiado o sensible.

Su aplicación es clave en casos donde se requiere extraer información visual de manera automatizada y con alta precisión, como en monitoreo de seguridad, validación de identidad, clasificación de archivos multimedia o análisis de evidencia audiovisual. Además, al integrarse con otros flujos de análisis, permite enriquecer los procesos de toma de decisiones con datos derivados de contenido visual, optimizando operaciones en sectores como justicia, salud, retail o servicios públicos. Esta funcionalidad contribuye a reducir la intervención manual, aumentar la trazabilidad y mejorar la eficiencia en entornos donde la imagen y el video son fuentes clave de información.

Modelado Predictivo y Recomendación

- **Amazon Forecast**: Este componente, integrado en el módulo de Analítica Avanzada de Houndoc, permite generar proyecciones precisas a partir de series temporales utilizando modelos de inteligencia artificial entrenados sobre datos históricos. Automatiza el proceso de modelado estadístico y aprendizaje automático para anticipar comportamientos futuros, tales como la demanda de productos, niveles de inventario, ingresos esperados o variaciones operativas.

Su capacidad para incorporar variables externas -como tendencias de mercado, calendarios, estacionalidades u otros factores contextuales- mejora significativamente la precisión de las predicciones, incluso en escenarios complejos o con alta volatilidad.

© Blend 2026. Todos los derechos reservados.

Esta funcionalidad es especialmente útil para optimizar la planificación estratégica, reducir el desperdicio, prevenir riesgos y facilitar la toma de decisiones basadas en evidencia. Su integración dentro de flujos de trabajo analíticos permite a las organizaciones anticiparse con mayor claridad y actuar con base en escenarios proyectados de forma automatizada y escalable.

- **Amazon Personalize**: Este componente, dentro del módulo de Analítica Avanzada de Houndoc, permite crear sistemas de recomendación personalizados utilizando inteligencia artificial. A partir del análisis de datos históricos sobre comportamiento de usuarios, preferencias y contexto de uso, genera sugerencias ajustadas a cada individuo, mejorando significativamente la relevancia de los contenidos o productos ofrecidos.

Esta funcionalidad permite aplicar modelos avanzados de machine learning sin necesidad de experiencia especializada en ciencia de datos, lo que facilita su adopción en entornos donde la personalización es clave: desde plataformas educativas o portales de servicios hasta motores de recomendación de contenido, productos financieros, bienes de consumo o estrategias de comunicación. La personalización generada contribuye a aumentar la interacción, mejorar la experiencia del usuario y fortalecer los resultados operativos, al entregar valor de forma más precisa y oportuna.

Entrenamiento y Despliegue de Modelos Propios

- **Amazon SageMaker**: Este componente, fundamental dentro del módulo de Analítica Avanzada de Houndoc, permite desarrollar, entrenar, optimizar e implementar modelos de inteligencia artificial a escala, facilitando la gestión completa del ciclo de vida de modelos de machine learning. Ofrece un entorno integrado que permite a científicos de datos e ingenieros trabajar de forma colaborativa y eficiente, reduciendo significativamente el tiempo y esfuerzo requeridos para poner en producción soluciones basadas en datos.

Es compatible con múltiples frameworks y fabricantes de modelos líderes en la industria, como **Hugging Face**, **TensorFlow**, **PyTorch** o **Scikit-learn**, lo que permite implementar modelos preentrenados o personalizados para tareas como clasificación, regresión, segmentación, predicción de eventos, detección de anomalías, entre otros. Por ejemplo, puede desplegar modelos de lenguaje como **BERT** para análisis semántico avanzado, o modelos de series temporales como **DeepAR** para previsión de demanda. También admite el ajuste fino de modelos fundacionales en contextos específicos, aumentando su precisión sin necesidad de desarrollos desde cero.

Además, su escalabilidad y capacidad de integración lo hacen ideal para entornos empresariales que requieren alta disponibilidad, auditoría, control de versiones y gobernanza sobre modelos, aportando un marco robusto para la automatización analítica y la toma de decisiones basada en modelos predictivos y prescriptivos.

Uso en migración. Analiza métricas de consumo y patrones de comunicación, y extrae texto, estructura y clasificación de repositorios documentales y multimedia durante la migración de contenido.

3.8. Dashboards

El componente **Dashboards** de Houndoc proporciona una capa de presentación analítica orientada a la toma de decisiones basada en datos. Su propósito es ofrecer a

© Blend 2026. Todos los derechos reservados.

los usuarios finales una experiencia visual clara, interactiva y contextualizada para monitorear indicadores clave, explorar tendencias, validar hipótesis y visualizar resultados de análisis avanzados.

Está diseñado para integrarse de forma nativa con la infraestructura analítica de la plataforma y soportar múltiples motores de visualización embebida, permitiendo consolidar la información procesada en tableros personalizables, seguros y de alto rendimiento.

Infraestructura Técnica y Servicios Asociados

- **Amazon QuickSight**

Es el motor de visualización principal en Houndoc, utilizado para construir dashboards interactivos con conectividad directa a fuentes como Amazon S3, Redshift, Athena, RDS o archivos tabulares. QuickSight permite desarrollar visualizaciones optimizadas para entornos web y móviles, con bajo tiempo de carga y capacidades de análisis visual enriquecido.

- **Amazon Q for QuickSight**

A través de Amazon Q, Houndoc habilita capacidades de análisis conversacional directamente sobre los dashboards. Esta funcionalidad permite a los usuarios realizar preguntas en lenguaje natural como *"¿Cuál fue la variación mensual del ingreso en la región sur?"*, y recibir respuestas automáticas en forma de gráficos, KPIs o explicaciones contextuales, sin necesidad de conocimientos técnicos en SQL o diseño de visualizaciones.

- **Soporte para herramientas externas (Power BI, Tableau)**

Para organizaciones que utilizan ecosistemas mixtos de visualización, el módulo permite integrar dashboards provenientes de otras plataformas como Power BI o Tableau. Estas integraciones se realizan mediante visualización embebida y autenticación federada, permitiendo mantener una experiencia unificada para el usuario final.

Capacidades Técnicas del Módulo

- **Visualización Interactiva y Navegación Contextual**

Los tableros permiten aplicar filtros dinámicos, selección cruzada, jerarquías de navegación (drill-down/drill-up), segmentaciones por usuario y exploración detallada por variable o métrica.

- **Diseño Responsive y Embebido**

Las visualizaciones pueden ser integradas dentro de módulos funcionales de Houndoc o en portales externos. El diseño es completamente adaptable a resoluciones móviles, de escritorio y pantallas de monitoreo.

- **Control de Acceso y Segmentación de Datos**

Cada dashboard puede restringir la información visible según el perfil del usuario, su rol o su dominio de datos. Esto garantiza cumplimiento con políticas de seguridad y confidencialidad, al tiempo que se mantiene la personalización de la experiencia.

- **Automatización y Distribución de Reportes**

© Blend 2026. Todos los derechos reservados.

Permite agendar la actualización de datos, distribución automática de reportes en PDF o Excel, y generar alertas visuales o por correo cuando se superan umbrales críticos.

- **Soporte Multifuente y Datos Derivados**

Cada visualización puede ser alimentada desde múltiples orígenes simultáneamente (datasets procesados, resultados de modelos, agregados desde data lake o queries en tiempo real), permitiendo integrar análisis predictivos, operativos y estratégicos en un mismo espacio visual.

Este componente es esencial para transformar los resultados analíticos y operativos en representaciones comprensibles y accionables, facilitando la democratización del análisis de datos en todos los niveles de la organización.

Uso en migración. Muestra el avance contra el programa, el estado de cada ola, las validaciones completadas y las incidencias.

4. Capacidades de datos, analítica e IA

Houndoc es una solución tecnológica de nueva generación orientada a potenciar las capacidades de análisis, automatización y toma de decisiones de organizaciones que gestionan grandes volúmenes de información heterogénea. Su propósito fundamental es transformar datos dispersos, no estructurados o difíciles de procesar en activos de información útiles, estructurados y accesibles, facilitando su explotación a través de herramientas de inteligencia artificial (IA), automatización de flujos de trabajo, motores de búsqueda y sistemas de visualización analítica.

En un entorno donde la disponibilidad de datos crece de forma exponencial, muchas instituciones enfrentan desafíos relacionados con la integración de fuentes, la validación y enriquecimiento de información, y la generación de conocimiento accionable. Houndoc resuelve estos problemas mediante un enfoque modular y escalable que permite consolidar repositorios de información, aplicar técnicas de procesamiento cognitivo, y entregar resultados comprensibles, trazables y auditables a través de interfaces accesibles para distintos perfiles de usuario.

El objetivo principal de la solución es acelerar los ciclos de procesamiento y análisis de datos, mejorar la calidad y trazabilidad de la información disponible, automatizar tareas manuales asociadas a la gestión de datos, y proporcionar herramientas avanzadas para el descubrimiento de patrones, relaciones y hallazgos relevantes en contextos operativos, regulatorios o investigativos.

Los beneficios clave para los usuarios incluyen:

- Disminución drástica de los tiempos requeridos para la búsqueda y consolidación de información.
- Mejora en la precisión de los análisis mediante la aplicación de modelos de IA y procesamiento semántico.
- Reducción de la carga operativa en tareas rutinarias a través de la automatización inteligente.

© Blend 2026. Todos los derechos reservados.

- Incremento en la capacidad institucional para responder de manera oportuna, basada en evidencia y con trazabilidad.

Houndoc está diseñado para atender a un público objetivo diverso, incluyendo usuarios operativos que requieren búsquedas rápidas y efectivas, analistas que necesitan realizar exploraciones profundas sobre grandes conjuntos de datos, tomadores de decisiones que exigen información consolidada en tiempo real, y equipos técnicos que buscan soluciones robustas, seguras y adaptables para entornos complejos.

4.1. SMA - Subscription Management

El componente **Subscription Management (SMA)** de Houndoc está diseñado para ofrecer una gestión robusta y flexible del ciclo de vida de suscripciones, pagos y administración financiera de usuarios y organizaciones. Está construido sobre una arquitectura especializada para la operación de modelos comerciales basados en suscripción, facturación por uso, y esquemas híbridos de cobro recurrente.

SMA está compuesto por tres submódulos principales que trabajan de forma coordinada para brindar una solución completa de administración de servicios bajo demanda:

Subscription Billing

Este submódulo proporciona control total sobre la gestión de planes de suscripción y estructuras tarifarias. Permite la definición de modelos de precios dinámicos que incluyen tarifas fijas, basadas en consumo o escalonadas. Soporta eventos como alta, baja, renovación, cancelación o suspensión de suscripciones, permitiendo una trazabilidad completa del historial financiero de cada cliente. También ofrece soporte para ciclos de facturación complejos, prorrateo, gestión de créditos y generación de facturas personalizadas.

Payment Platform

SMA se integra fácilmente con plataformas de pago externas, bancos emisores y herramientas antifraude mediante un ecosistema de plugins extensibles. Esta capa permite gestionar transacciones de pago, emisión de órdenes de cobro, conciliación y validación de pagos en tiempo real. Su diseño orientado a extensibilidad permite a cada organización adaptar la integración con sus propios proveedores financieros y métodos de pago.

Admin User Interface

El componente incluye una interfaz administrativa web (back-office) diseñada para equipos de finanzas, operaciones, atención al cliente y desarrollo. Esta interfaz permite gestionar planes de precios, revisar el estado de las suscripciones activas, modificar parámetros contractuales y visualizar el historial de transacciones. Está orientada a brindar autonomía a los equipos funcionales, con un enfoque en la usabilidad, segmentación por rol y control seguro de cambios.

SMA es un pilar clave en proyectos donde la monetización de capacidades tecnológicas o el control operativo de clientes y organizaciones usuarias es un requerimiento estratégico. Su arquitectura desacoplada y orientada a eventos permite

© Blend 2026. Todos los derechos reservados.

su integración con otros módulos de Houndoc y sistemas externos de gestión comercial o financiera.

4.2. MAN - Maps Analytics

El componente **Maps Analytics (MAN)** de Houndoc está diseñado para el análisis geoespacial avanzado, facilitando la visualización, segmentación y exploración de datos geoespaciales. Su objetivo principal es permitir a los usuarios identificar patrones, correlaciones y fenómenos en función de su ubicación, mediante una interfaz visual altamente interactiva y personalizable.

Este módulo es especialmente útil en casos de uso donde la variable geográfica es un eje de análisis clave, como en investigaciones territoriales, monitoreo ambiental, análisis de movilidad, distribución de eventos, y evaluación de cobertura de servicios.

Visualización Geoespacial Dinámica

MAN permite representar información en múltiples niveles de detalle utilizando capas de datos vectoriales y ráster, integrando diferentes tipos de visualización como puntos, polígonos, rutas, mapas de calor y capas categóricas. La plataforma soporta múltiples proyecciones cartográficas y escalas configurables, lo cual permite combinar datos provenientes de diversas fuentes geográficas.

El usuario puede aplicar filtros sobre atributos no espaciales (por ejemplo: fecha, tipo de evento, entidad involucrada) y obtener representaciones dinámicas que cambian en tiempo real, facilitando la exploración desde distintas perspectivas.

Integración con Datos Internos y Externos

MAN permite importar y consumir fuentes geográficas internas de la organización (bases de datos espaciales, archivos shapefile, CSV con coordenadas, etc.) y externas como servicios web de mapas, APIs públicas o datos abiertos en formatos geoespaciales estándar (GeoJSON, WKT, KML). Esta capacidad facilita la integración de datos operativos con información contextual relevante.

También permite vincular eventos o entidades dentro de Houndoc a coordenadas geográficas o zonas administrativas, generando representaciones automáticas de casos, alertas o registros localizados.

Análisis Espacial y Comparativo

El módulo ofrece herramientas analíticas como:

- Cálculo de densidades, concentraciones y cobertura geográfica.
- Comparación entre periodos o grupos de datos por región.
- Detección de zonas de calor y generación de superficies de influencia.

Estas capacidades permiten transformar datos crudos en conocimiento espacialmente distribuido, útil para detectar riesgos, planificar intervenciones o identificar áreas críticas.

© Blend 2026. Todos los derechos reservados.

Interfaz y Experiencia de Usuario

La interfaz de MAN está diseñada para facilitar la interacción de usuarios tanto técnicos como no técnicos. Incluye herramientas de navegación como zoom, filtros temporales, selección de capas, leyendas interactivas y paneles de detalle que permiten visualizar los atributos asociados a los elementos geográficos.

El componente también ofrece soporte para exportación de mapas en distintos formatos (PNG, SVG, GeoJSON) y generación de informes automáticos con gráficos vinculados a las visualizaciones geográficas.

Componentes Técnicos

Arquitectura General con AWS

- **Almacenamiento de Datos:**

+ **Amazon S3**: Se utilizará para almacenar datos geoespaciales en formatos como shapefiles, GeoJSON, KML, y otros archivos CSV con coordenadas. S3 servirá como el repositorio principal para estos datos antes de su procesamiento y visualización. + **Amazon RDS con PostgreSQL**: El servicio RDS alojará bases de datos espaciales, utilizando PostgreSQL con la extensión PostGIS para almacenar y gestionar información geoespacial en formato estructurado, como puntos, polígonos, y rutas. Este servicio es ideal para manejar consultas espaciales avanzadas y proporcionar los datos al módulo de visualización.

- **Procesamiento y Análisis:**

+ **AWS Lambda**: Usado para ejecutar código sin necesidad de aprovisionar servidores. Lambda puede procesar y transformar los datos geoespaciales almacenados en S3 o RDS antes de pasarlos al frontend para su visualización. También puede manejar tareas como la carga de datos en tiempo real o la ejecución de análisis espaciales, como el cálculo de densidades o la comparación de datos entre periodos. + **Amazon Secrets Manager**: Se utiliza para gestionar credenciales de acceso y otras variables sensibles, asegurando que las conexiones a RDS, S3 y otros servicios de AWS sean seguras y gestionadas correctamente.

- **Redes y Aprovisionamiento:**

+ **Amazon VPC**: Una VPC privada se configura para asegurar que los servicios, como RDS y Lambda, estén aislados y protegidos. El tráfico entre servicios dentro de la VPC será privado, y se utilizarán subredes adecuadas para garantizar la seguridad de los datos. + **Amazon Route 53**: Para gestionar el enrutamiento DNS, garantizando que las solicitudes externas a la plataforma de visualización geoespacial lleguen correctamente a las instancias de backend en ECS.

- **Aprovisionamiento del Backend:**

+ **Amazon ECS y Fargate**: ECS con Fargate será utilizado para gestionar el backend que sirve la plataforma de visualización. Fargate facilita la ejecución de contenedores sin necesidad de gestionar la infraestructura subyacente, permitiendo escalar dinámicamente según la carga.

- **Interfaz de Usuario:**

© Blend 2026. Todos los derechos reservados.

+ **Kepler.gl**: Se usará como herramienta externa para la visualización geoespacial. Kepler.gl es una librería de JavaScript que se puede integrar con los datos geoespaciales provenientes de S3 y RDS, ofreciendo capacidades avanzadas como visualización en múltiples capas, mapas de calor, y análisis comparativo. + **Frontend con librerías de lenguaje**: Se integrarán librerías personalizadas de JavaScript o frameworks front-end (como React o Angular) para construir la interfaz de usuario y mostrar las visualizaciones dinámicas. Esto incluirá funcionalidades como la navegación por el mapa, la aplicación de filtros, y la exportación de mapas en diferentes formatos (PNG, SVG, GeoJSON).

Flujo de Datos

- **Importación de Datos**: Los datos geoespaciales (shapefiles, CSV, GeoJSON) se almacenan en **S3** o en **RDS**. Estos pueden provenir tanto de fuentes internas (bases de datos de la organización) como externas (APIs públicas, datos abiertos).
- **Procesamiento**: **AWS Lambda** procesa estos datos, realiza análisis espaciales (como el cálculo de densidades, zonas de calor, etc.), y los prepara para ser visualizados.
- **Visualización**: **Kepler.gl** se encarga de renderizar los datos en el frontend. Los usuarios pueden aplicar filtros, ver diferentes capas de información, y explorar los datos de manera dinámica.
- **Exportación**: Los usuarios pueden exportar las visualizaciones generadas como imágenes (PNG, SVG) o datos (GeoJSON) para su posterior análisis o informes.

Capacidades Analíticas

- **Cálculo de Densidades y Cobertura**: Lambda puede ejecutar consultas espaciales en RDS (PostGIS) para calcular áreas de alta concentración o cobertura geográfica.
- **Comparación entre Periodos**: Los datos almacenados en RDS pueden ser analizados para comparar diferentes periodos o grupos de datos por región, generando informes comparativos.
- **Generación de Superficies de Influencia**: Lambda y RDS trabajarán juntos para calcular áreas de influencia o zonas de calor, los cuales pueden ser visualizados en Kepler.gl.

Seguridad

- **IAM Roles**: Se utilizarán roles de IAM para gestionar los permisos y accesos entre los servicios de AWS (Lambda, RDS, S3, etc.).
- **Amazon Secrets Manager**: Asegura que las credenciales de la base de datos y otros accesos sensibles se mantengan seguros.

4.3. BAO - Business Automation and Optimization

El componente **Business Automation and Optimization (BAO)** de Houndoc está diseñado para modelar, ejecutar y optimizar procesos de negocio complejos de manera automatizada. Actúa como el motor de lógica operativa de la plataforma,

© Blend 2026. Todos los derechos reservados.

permitiendo implementar reglas dinámicas, flujos condicionados, y mecanismos de toma de decisiones automatizada en entornos variables y multientidad.

BAO facilita la digitalización de procesos mediante el uso de motores de reglas declarativas y sistemas de planificación basados en restricciones. Esta capacidad permite responder de forma flexible a escenarios cambiantes, garantizar la coherencia en la ejecución de políticas, y reducir la intervención manual en tareas repetitivas o sujetas a validaciones estructuradas.

Arquitectura del Módulo

BAO se compone de tres capas técnicas principales:

- **Motor de Reglas de Negocio**

Implementa un sistema basado en reglas declarativas que permite definir condiciones, excepciones, dependencias y consecuencias para decisiones automatizadas. Las reglas son interpretadas en tiempo de ejecución, lo que permite actualizar la lógica sin necesidad de modificar código fuente. Este motor se integra con otros módulos de Houndoc, permitiendo aplicar lógica compleja sobre datos analizados, resultados de modelos o entradas de formularios.

- **Motor de Ejecución de Procesos**

Utiliza una arquitectura basada en flujos BPMN (Business Process Model and Notation) que permite definir y desplegar procesos con lógica secuencial, paralela o asincrónica. Incluye componentes para la ejecución de tareas humanas, llamadas a servicios externos, envío de notificaciones y manejo de excepciones. Los procesos pueden ser monitoreados en tiempo real y auditados para trazabilidad completa.

- **Motor de Optimización**

Permite resolver problemas de asignación, planificación y optimización multiobjetivo. Emplea algoritmos heurísticos y planificación basada en restricciones para identificar soluciones óptimas bajo reglas definidas, por ejemplo: asignación de recursos, calendarización de actividades, o generación de planes operativos ajustados a criterios de eficiencia.

Integración con la Plataforma

BAO se comunica de forma nativa con otros componentes de Houndoc, permitiendo que decisiones derivadas de análisis de datos (AAN), resultados de formularios (AIF), o eventos provenientes de flujos (WFM), activen procesos automatizados bajo reglas configurables.

Además, sus interfaces están diseñadas para ser operadas tanto por usuarios técnicos como por equipos de negocio, permitiendo que analistas funcionales definan reglas sin necesidad de intervención directa del equipo de desarrollo, facilitando la gobernanza operativa.

4.4. AIS - AI Search Engine

El componente **AI Search Engine (AIS)** de Houndoc proporciona un sistema de búsqueda avanzada diseñado para ofrecer acceso rápido, preciso y semánticamente relevante a grandes volúmenes de información estructurada y no estructurada. Su

© Blend 2026. Todos los derechos reservados.

arquitectura está optimizada para soportar cargas de trabajo de búsqueda intensiva en entornos empresariales, regulados y de misión crítica.

AIS está construido sobre **Amazon OpenSearch Service**, la solución administrada en la nube de AWS basada en Elasticsearch, lo que garantiza escalabilidad, alta disponibilidad, cifrado de extremo a extremo y rendimiento constante en ambientes de producción.

Capacidades Técnicas Clave

- **Indexación de Texto Completo**

Utiliza esquemas de tokenización y análisis lingüístico avanzados para habilitar búsquedas por palabras clave, frases exactas, coincidencia parcial, sinónimos y variaciones gramaticales. Esto permite una recuperación de información más flexible y robusta que la simple coincidencia literal.

- **Análisis Semántico y Filtrado Contextual**

AIS integra capacidades de análisis morfosintáctico que permiten enriquecer los contenidos indexados con metadatos, entidades extraídas, temas, fechas y estructuras jerárquicas. Esto habilita búsquedas contextuales y segmentación dinámica por tipo de documento, dominio temático o etapa del proceso.

Adicionalmente, permite configurar categorías jerárquicas personalizadas para facilitar la organización temática de los resultados y mejorar la navegación visual en la interfaz de usuario.

- **Consultas Compuestas y Navegación por Facetas**

Permite construir consultas que combinan texto libre con condiciones estructuradas sobre múltiples campos, fechas, rangos numéricos o atributos personalizados. La navegación por facetas permite explorar el contenido desde filtros derivados automáticamente del índice.

La configuración de filtros admite múltiples tipos de campos, tales como texto, fechas, listas seleccionables y valores múltiples, lo cual ofrece mayor flexibilidad al momento de refinar las búsquedas dentro del sistema.

- **Relevancia y Personalización de Resultados**

Gracias a los mecanismos de rankeo configurables de OpenSearch, AIS puede ajustar la puntuación de los resultados según reglas de negocio, patrones de consulta históricos o señales contextuales. También puede integrarse con modelos de aprendizaje supervisado para personalizar resultados a usuarios o perfiles específicos.

- **Indexación en Tiempo Real y Alta Disponibilidad**

Nuevos datos o documentos pueden ser indexados y consultados en cuestión de segundos gracias al soporte de ingestión en streaming y replicación automática de OpenSearch. La infraestructura subyacente soporta despliegues multi-nodo y balanceo de carga para entornos exigentes.

Esto permite mantener la disponibilidad del servicio incluso bajo alta demanda, con capacidad de recuperación rápida ante fallos o picos de tráfico.

- **Interfaces de Consulta Accesibles**

© Blend 2026. Todos los derechos reservados.

AIS ofrece tanto una interfaz de usuario basada en web para consultas interactivas, como endpoints REST para integrarse con otros módulos o sistemas externos. Soporta autocompletado, sugerencias, búsquedas tolerantes a errores (fuzzy), filtros avanzados y exportación de resultados.

Asimismo, proporciona capacidades de visualización directa para documentos en diversos formatos, incluyendo archivos de texto, hojas de cálculo, audios, videos, imágenes y datos geoespaciales, todos accesibles desde almacenamiento S3. Formatos soportados: PDF, HTML, DOCX, XLSX, imágenes, audio, video, GeoJSON y SHP.

Seguridad y Auditoría

AIS hereda los mecanismos de seguridad gestionada de Amazon OpenSearch Service, incluyendo:

- Autenticación basada en IAM o SSO.
- Cifrado en tránsito (TLS) y en reposo.
- Control de acceso granular por índice, tipo de dato o acción (búsqueda, escritura, administración).
- Auditoría de accesos y consultas.

Las credenciales de acceso al motor de búsqueda y sus endpoints se administran de manera segura mediante AWS Secrets Manager, con integración nativa a servicios de autenticación como Keycloak vía OAuth2.

Este componente es fundamental en contextos donde se requiere navegar rápidamente por información crítica como expedientes jurídicos, documentos normativos, historiales clínicos, bases de conocimiento o registros operativos.

Configuración y Despliegue

AIS permite una configuración flexible y modular de los índices a través de parámetros como `index_search`, `index_filter`, `index_category`, `index_dataMain` e `index_label`, los cuales definen qué campos son visualizados, filtrados o clasificados dentro de la plataforma. Cada índice debe tener su propia configuración y mapeo de campos.

El despliegue de la solución se realiza mediante AWS CDK y CodePipeline, habilitando infraestructura como código para los stacks de frontend y backend, los cuales operan sobre ECS Fargate con configuración de redes privadas, balanceo de carga y versiones controladas del sistema. Los recursos necesarios incluyen buckets de almacenamiento en S3, roles IAM, clientes de autenticación en servidores IAM, y configuración de entornos a través de archivos `.env` y `version_config.json`.

Arquitectura según el entorno de despliegue

El componente AIS ha sido diseñado para operar tanto en entornos cloud como on-premises, adaptándose a distintos modelos operativos y niveles de autonomía técnica según los requerimientos de cada organización. En implementaciones en la nube, se utiliza Amazon OpenSearch Service, una solución administrada y compatible con Elasticsearch que permite aprovisionar, escalar y asegurar el servicio sin necesidad de gestionar manualmente los nodos del clúster.

© Blend 2026. Todos los derechos reservados.

Por su parte, en entornos on-premises, AIS se basa directamente en Elasticsearch, desplegado sobre infraestructura local bajo la administración directa del equipo técnico de la organización. En este modelo, la operación del clúster requiere aprovisionamiento de recursos físicos o virtuales, configuración de red, políticas de seguridad, manejo de réplicas y escalado horizontal, todo gestionado internamente.

Esta diferenciación permite que AIS opere tanto en arquitecturas modernas basadas en servicios administrados como en entornos con políticas de control estricto, soberanía de datos o integración con ecosistemas corporativos cerrados.

Funcionamiento de Elasticsearch: fundamentos técnicos

Elasticsearch constituye la base tecnológica sobre la que opera AIS, tanto en su versión administrada (OpenSearch) como en entornos controlados directamente por la organización. Es un motor de búsqueda distribuido que organiza la información mediante un sistema de índices, shards (fragmentos) y réplicas, lo que permite distribuir la carga de trabajo, asegurar la disponibilidad y escalar horizontalmente con eficiencia.

Durante la indexación, los documentos son analizados mediante técnicas de procesamiento lingüístico: tokenización, eliminación de términos irrelevantes (stop words), y normalización. Estos procesos generan estructuras invertidas que permiten búsquedas de alta velocidad, con soporte para coincidencias exactas, texto libre, operadores booleanos y filtros estructurados.

Las consultas se definen mediante el lenguaje Query DSL, que permite expresar condiciones complejas utilizando operadores como match, range, term, bool, must, should y filter. Esta flexibilidad permite implementar reglas de negocio dentro del motor de búsqueda y personalizar la relevancia de los resultados mediante puntajes ajustables.

Además, Elasticsearch soporta agregaciones, una capacidad avanzada que permite realizar estadísticas, análisis de tendencias o reportes directamente sobre los datos indexados, sin necesidad de extraerlos a herramientas externas.

Aplicaciones Típicas

Este componente es fundamental en contextos donde se requiere navegar rápidamente por información crítica como expedientes jurídicos, documentos normativos, historiales clínicos, bases de conocimiento o registros operativos.

Asimismo, se adapta a entornos donde se requiere gestionar contenido multimedia, visualización de mapas judiciales o consolidación de repositorios documentales con formatos mixtos.

4.5. WebScrapping

El componente **WebScrapping** de Houndoc permite la recolección automatizada de información desde fuentes web públicas o privadas, estructuradas y no estructuradas, como parte de los procesos de enriquecimiento de datos, vigilancia estratégica, monitoreo de medios o verificación de contenido. Este módulo actúa como un canal de adquisición de datos externos que pueden ser validados, transformados y procesados dentro de los flujos internos de la plataforma.

© Blend 2026. Todos los derechos reservados.

Diseñado con una **arquitectura flexible y desacoplada**, WebScrapping se adapta dinámicamente a distintos tipos de portales, estructuras HTML, APIs abiertas o repositorios no indexados, ofreciendo una solución resiliente a cambios frecuentes en las interfaces web.

Arquitectura de Referencia sobre AWS

El componente WebScrapping de Houndoc se soporta en una arquitectura escalable y desacoplada implementada sobre **servicios nativos de AWS**, lo cual permite una operación eficiente, segura y de bajo mantenimiento.

Esta arquitectura se compone de las siguientes capas funcionales:

1. Capa de Ingestión

Responsable de obtener los enlaces de las fuentes objetivo (por ejemplo, artículos de noticias, boletines o comunicados).

- **AWS Lambda** se encarga de ejecutar funciones que recorren las fuentes para identificar nuevos enlaces.
- **AWS Glue** se emplea para transformar y extraer metadatos relevantes.
- **AWS Step Functions** orquesta las tareas de scraping de manera secuencial y controlada.

2. Capa de Login

Gestiona credenciales y accesos a portales protegidos o privados.

- **AWS Lambda** obtiene las credenciales dinámicamente.
- **Amazon DynamoDB** almacena de forma segura las credenciales y tokens de sesión.

3. Capa de Scraping

Realiza la obtención del contenido específico a partir de los enlaces previamente detectados.

- **AWS Lambda** ejecuta el scraping de contenido.
- **AWS Glue** limpia y transforma la información extraída.
- **Amazon S3** almacena el contenido crudo y limpio de forma segura y escalable.

4. Capa de Clasificación

Aplica algoritmos de aprendizaje automático para categorizar los contenidos obtenidos.

- **Amazon EventBridge** coordina eventos disparadores del proceso.
- **Amazon SageMaker** entrena y ejecuta modelos de clasificación.
- **Amazon S3** almacena los resultados y las clasificaciones obtenidas.

5. Capa de Explotación de Datos

Permite la consulta, análisis y visualización del contenido extraído y clasificado.

- **Amazon Athena** permite consultas SQL sobre los datos almacenados en S3.

© Blend 2026. Todos los derechos reservados.

- **Amazon QuickSight** genera tableros e informes visuales de consumo para usuarios finales.

Componentes Técnicos

- **Orquestador de Tareas de Extracción**

Este subcomponente define y ejecuta scripts de extracción para cada fuente configurada, aplicando técnicas de scraping adaptativo y detección estructural. La lógica de extracción puede ejecutarse en intervalos programados o mediante disparadores por eventos, integrándose con el módulo WFM.

- **Agentes de Scraping**

El módulo soporta múltiples motores de scraping configurables por tipo de fuente, entre los cuales se pueden incluir herramientas basadas en navegación (headless browsers), extracción por XPath/CSS selectors o consumo directo de endpoints JSON. Cada agente puede ser configurado con headers, autenticación, cookies y lógica personalizada para manejar paginación, formularios dinámicos o AJAX.

- **Limpieza y Normalización de Datos**

Una vez obtenida la información, los datos pasan por un proceso de depuración para eliminar ruido, validar formatos, extraer entidades clave y convertir el contenido a estructuras homogéneas (por ejemplo, JSON normalizado, tabular, etc.). Esto permite su integración con el catálogo, flujos de análisis o modelos de IA.

- **Verificación y Control de Calidad**

El componente puede aplicar filtros para verificar la confiabilidad de las fuentes, validar la consistencia del contenido, detectar duplicidades y aplicar reglas de control semántico sobre los textos extraídos. También permite priorizar dominios certificados o institucionales en función de criterios de confianza definidos por el cliente.

- **Listas Blancas, Grises y Negras**

El sistema incorpora un mecanismo de clasificación de fuentes mediante listas blancas, grises y negras, que permite al cliente definir el nivel de confianza de cada fuente de información. Las listas blancas agrupan fuentes altamente confiables y priorizadas en los procesos de extracción; las listas grises contienen fuentes parcialmente confiables que pueden requerir validación adicional; y las listas negras excluyen fuentes no confiables, de las cuales no se extrae información bajo ninguna circunstancia. Esta clasificación se aplica de forma dinámica durante la ejecución de las tareas de scraping y verificación.

- **Persistencia y Auditoría**

Todo el contenido extraído es almacenado en un repositorio de datos intermedio, etiquetado con su origen, fecha de recolección y estado de validación. Este repositorio puede ser explorado, depurado o auditado por equipos técnicos para control de calidad o resolución de incidencias.

4.6. AIF - Automated Intelligent Forms

El componente **AIF (Automated Intelligent Forms)** de Houndoc está orientado a la generación, despliegue y procesamiento automatizado de formularios digitales

© Blend 2026. Todos los derechos reservados.

dinámicos. Su propósito es facilitar la recolección estructurada de información en procesos institucionales, legales, administrativos o técnicos, permitiendo una definición declarativa de lógica condicional, validaciones de campo, auto-población de valores y automatización documental.

AIF permite a las organizaciones definir formularios complejos con estructura flexible, reglas de negocio embebidas y generación automática de documentos finales en formatos reutilizables.

Estructura Técnica del Componente

AIF se basa en una arquitectura declarativa compuesta por tres capas:

- **Definición de Estructura (YAML/Markdown)**

Los formularios son definidos utilizando estructuras YAML o Markdown enriquecidas, que describen campos, tipos de dato, jerarquías, dependencias lógicas, valores predeterminados y reglas de presentación. Esta abstracción permite modificar formularios sin necesidad de código, facilitando su mantenimiento por parte de usuarios funcionales.

- **Motor de Renderizado Dinámico**

Este motor interpreta la definición de cada formulario y genera interfaces interactivas adaptadas al tipo de usuario y al contexto del proceso. Soporta lógica condicional para mostrar u ocultar secciones, cálculos automáticos, despliegue progresivo de secciones y control de flujo entre pasos.

- **Procesamiento y Validación**

Una vez diligenciado, el formulario pasa por un proceso de validación automática que asegura la integridad de los datos recolectados. Las validaciones pueden ser de tipo estructural (obligatoriedad, formatos, rangos) o condicional, con base en reglas definidas. El resultado puede ser almacenado como estructura JSON, enviado a motores de decisión (como BAO), o transformado en un documento PDF, DOCX o HTML estandarizado.

Capacidades Adicionales

- **Plantillas Inteligentes**

Los formularios pueden incluir secciones prellenadas desde otras fuentes (por ejemplo, datos de usuario, sistemas externos o resultados de análisis previos) y guardar plantillas reutilizables para casos frecuentes o masivos.

- **Versionamiento y Reutilización**

AIF soporta múltiples versiones de un mismo formulario, permitiendo mantener trazabilidad de los cambios y adaptar los formularios a nuevas normativas sin afectar formularios ya diligenciados.

- **Exportación y Firma Digital**

Los formularios pueden convertirse automáticamente en documentos finales listos para firma, archivo o notificación, integrándose con módulos de gestión documental o plataformas de firma electrónica.

- **Integración con flujos y motores de reglas**

© Blend 2026. Todos los derechos reservados.

AIF se comunica nativamente con **WFM** para desencadenar acciones al momento de la recepción del formulario, y con **BAO** para aplicar lógica automatizada a los datos ingresados.

4.7. NSB - Nadia Smart Buddy

El componente **Nadia Smart Buddy (NSB)** es el asistente cognitivo de Houndoc, diseñado para interactuar con usuarios finales mediante lenguaje natural, a través de múltiples canales como interfaces web, mensajería instantánea, voz o integraciones embebidas en portales corporativos. NSB permite automatizar la atención, el acompañamiento y la navegación contextualizada dentro de sistemas complejos, ofreciendo respuestas precisas, guiadas y contextualizadas a partir de bases de conocimiento estructuradas y no estructuradas.

A diferencia de otros módulos orientados a procesamiento backend, NSB está centrado en la experiencia de usuario final, permitiendo la configuración de agentes conversacionales adaptados a distintos dominios funcionales como salud, legal, educación, atención ciudadana o servicios internos.

Arquitectura Técnica y Capacidades

- **Modelo basado en Retrieval-Augmented Generation (RAG)**

NSB utiliza una arquitectura híbrida de generación con recuperación, en la que el modelo de lenguaje genera respuestas basadas en contexto, apoyado por un sistema de recuperación semántica de información (vector search). Esta estrategia permite combinar precisión factual con fluidez conversacional, manteniendo alineación con fuentes autorizadas.

- **Motor de diálogo con control de contexto**

Soporta conversaciones persistentes y sesiones multietapa, manteniendo el estado de diálogo y permitiendo navegar procesos complejos con lógica condicional, validación de entradas, confirmaciones, recomendaciones y acciones derivadas.

- **Motor de recuperación semántica**

Las respuestas generadas por NSB se apoyan en búsquedas semánticas sobre bases documentales estructuradas y no estructuradas. Este motor utiliza vectores embebidos generados a partir de los documentos indexados, permitiendo responder preguntas que requieren comprensión contextual, más allá de coincidencias literales.

- **Capa de configuración y entrenamiento**

NSB incluye un backoffice que permite parametrizar el comportamiento del agente: definición de dominio, tono, cobertura temática, inclusión de fuentes específicas, vocabulario clave y definición de intenciones. También admite correcciones manuales, feedback y curaduría de respuestas.

- **Control de seguridad y auditoría**

El módulo incluye mecanismos de autenticación (anónima, por token o federada), control de acceso por perfil, enmascaramiento de respuestas sensibles y registro de trazas de interacción para auditoría, análisis de satisfacción y mejora continua.

© Blend 2026. Todos los derechos reservados.

Integración con capacidades cognitivas

NSB puede invocar capacidades analíticas de otros módulos de Houndoc como extracción de texto, clasificación de documentos, análisis semántico o generación automatizada, para enriquecer sus respuestas. También puede consultar respuestas almacenadas, historial de interacciones previas o documentos institucionales para sostener su argumentación.

Este componente es ideal para escenarios en los que se requiere un agente inteligente que entienda, responda y actúe, mejorando la experiencia del usuario final con capacidades conversacionales avanzadas.

Implementación Multiplataforma

La **integración de WhatsApp API** permite automatizar la comunicación con los usuarios a través de múltiples canales, como webchat, WhatsApp, asistentes de voz (por ejemplo, integraciones con Contact Centers), correo electrónico o embebido en sistemas propietarios mediante APIs. Esta flexibilidad asegura que el sistema pueda adaptarse a las políticas de diseño y las necesidades específicas de cada cliente, permitiendo una experiencia de usuario personalizada y fluida.

Requiere una cuenta de WhatsApp Business vinculada a la Meta Business Suite, acceso a la API de WhatsApp y un servidor REST con SSL para manejar las solicitudes. La integración incluye la configuración de un **webhook** que recibe notificaciones en tiempo real sobre eventos como mensajes entrantes o cambios de estado, verificando su autenticidad mediante un token de verificación.

Además, la API permite el **envío de mensajes** automáticos a los usuarios, con un Access Token para autorizar la acción. Es esencial implementar un sistema de **monitoreo** para registrar interacciones y asegurar el buen funcionamiento de la integración. La seguridad también es clave, por lo que se debe manejar de manera segura el Access Token y controlar el acceso mediante el IAM, asignando roles específicos a los usuarios.

Por último, se deben realizar **pruebas exhaustivas** para garantizar que el sistema funcione correctamente, asegurando una comunicación eficiente y segura a través de los distintos canales. En resumen, esta integración ofrece una forma flexible y automatizada de interactuar con los usuarios, mejorando la atención al cliente y optimizando procesos comerciales.

Capacidades Técnicas

Autenticación

- **Amazon Cognito User Pool**: Gestiona la autenticación de usuarios, asegurando que los usuarios interactúan de manera segura con el sistema NSB a través de diferentes proveedores de identidad (como Google). Esto es clave para el acceso y la seguridad del usuario.
- **Frontend**:
- **AWS WAF**: Protege el frontend de amenazas web, asegurando la seguridad de las interacciones de los usuarios con la aplicación React alojada en S3.

© Blend 2026. Todos los derechos reservados.

- **Amazon CloudFront**: Entrega la aplicación React de manera eficiente a través de regiones, garantizando tiempos de carga rápidos para los usuarios.

API Backend

- **API Gateway**: Gestiona y enruta las solicitudes API a los diversos servicios dentro del backend, permitiendo la comunicación entre el frontend y los componentes del backend.
- **AWS Lambda (FastAPI)**: Ejecuta la lógica del backend, procesa las solicitudes e interactúa con los agentes conversacionales.

Respuesta en Streaming

- **API Gateway (WebSocket)**: Facilita la comunicación en tiempo real entre los usuarios y el backend, asegurando interacciones continuas y sin interrupciones para agentes de chat o de voz.
- **AWS Lambda**: Procesa interacciones en tiempo real, garantizando respuestas rápidas a las solicitudes de los usuarios.

LLM (Modelos de Lenguaje Grande)

- **Amazon Bedrock Converse API**: Potencia las capacidades de inteligencia conversacional del sistema, proporcionando el modelo de lenguaje necesario para generar respuestas significativas y precisas basadas en la entrada del usuario.

RAG Source (Recuperación Aumentada por Generación)

- **Amazon Bedrock Knowledge Bases**: Sirve como base para recuperar información relevante durante las conversaciones. Asegura que el sistema obtenga contenido autorizado y basado en hechos para generar respuestas.
- **OpenSearch Serverless**: Permite realizar búsquedas dentro de documentos o bases de conocimiento para proporcionar contexto que respalde la inteligencia conversacional y garantizar respuestas precisas.

Datos y Análisis

- **Amazon Athena**: Se utiliza para consultar los registros y datos almacenados en S3, permitiendo que los administradores analicen el uso y el rendimiento de los agentes conversacionales.
- **Amazon DynamoDB**: Almacena registros de conversaciones y datos de interacción con el bot, proporcionando la base de datos para gestionar los estados de las sesiones y los registros históricos de interacción.
- **EventBridge Pipes**: Integra datos de eventos en tiempo real y activa diferentes procesos, como alertar a los administradores o procesar retroalimentación de las interacciones de los usuarios.

Conocimiento Externo

- **S3 Bucket (Documentos)**: Almacena documentos y bases de conocimiento que el sistema consulta para generar respuestas precisas y contextualizadas.
- **Internet**: APIs o fuentes de datos externas que pueden mejorar las respuestas al integrar conocimientos basados en internet.

© Blend 2026. Todos los derechos reservados.

4.8. Scientific Notebooks

El componente **Scientific Notebooks** de Houndoc proporciona un entorno de trabajo interactivo para científicos de datos, analistas técnicos y desarrolladores especializados en procesamiento de información avanzada. Este módulo está diseñado para facilitar la exploración, validación, modelado y documentación de procesos analíticos complejos mediante cuadernos ejecutables que combinan código, visualizaciones y texto descriptivo en un solo flujo de trabajo.

Los notebooks permiten trabajar de manera reproducible y colaborativa sobre los datos gestionados por la plataforma, habilitando un entorno de experimentación que se integra con la infraestructura de datos, modelos y almacenamiento ya configurada en Houndoc.

Arquitectura Técnica

Scientific Notebooks se basa en la ejecución de cuadernos tipo **Jupyter Notebook o Zeppelin Notebook**, desplegados en entornos controlados y seguros dentro de la nube, con soporte para múltiples lenguajes como Python, R o SQL. Estos notebooks están conectados a los entornos de datos y procesamiento de Houndoc mediante capas de integración que permiten:

- **Conexión directa al data lake**: Permiten cargar y manipular datasets almacenados en Amazon S3, a través de interfaces programáticas o conectores compatibles con formatos Parquet, CSV, JSON y bases SQL.
- **Acceso a catálogos y metadatos**: Los notebooks pueden consultar el Glue Data Catalog para descubrir estructuras de datos, explorar esquemas, verificar versiones y entender el linaje de los datos utilizados en cada análisis.
- **Ejecución en entornos aislados**: Cada notebook es desplegado sobre un entorno aislado, autogestionado y escalable, que puede ser aprovisionado automáticamente al iniciarse una sesión. Estos entornos soportan librerías especializadas de procesamiento de datos, machine learning, visualización y estadística.
- **Persistencia y versionamiento de código**: Todo el contenido de los notebooks puede ser almacenado en repositorios vinculados con herramientas de versionamiento, lo que permite recuperar versiones anteriores, realizar revisiones o auditar experimentos pasados.

Capacidades Extendidas

- **Entrenamiento y prueba de modelos**: Los usuarios pueden construir, entrenar, evaluar y exportar modelos de machine learning dentro del entorno del notebook, integrando directamente con servicios como Amazon SageMaker para escalar el entrenamiento o desplegar modelos.
- **Automatización de tareas repetitivas**: Los notebooks pueden ser programados para ejecutarse de forma automática dentro de flujos definidos, por ejemplo, para generar reportes periódicos, actualizar indicadores o ejecutar pruebas de validación de datos.

© Blend 2026. Todos los derechos reservados.

- **Entorno colaborativo**: Los notebooks pueden ser compartidos entre usuarios con permisos diferenciados (lectura, edición, ejecución), permitiendo trabajo colaborativo y revisión por pares en proyectos de análisis complejos.

Este componente es clave para equipos técnicos que requieren flexibilidad y control total sobre el ciclo analítico, con acceso directo a la infraestructura y los datos, sin perder trazabilidad, seguridad ni conexión con el resto del ecosistema Houndoc.

4.9. Servicios complementarios en la nube o entorno local

Los servicios complementarios de Houndoc fortalecen su propuesta de valor al incorporar capacidades especializadas de procesamiento cognitivo, automatización documental, y transformación de datos no estructurados, todo sobre una arquitectura completamente escalable basada en servicios gestionados de **Amazon Web Services (AWS)**. Estos servicios operan como microcomponentes desacoplados y orquestables dentro de los flujos generales de la plataforma, habilitando soluciones inteligentes con bajo mantenimiento y alta eficiencia operativa.

4.9.1. Extracción de Texto y Reconocimiento Óptico de Caracteres (OCR)

Este servicio se apoya en **Amazon Textract**, que permite extraer texto impreso y manuscrito desde documentos escaneados o imágenes digitales. A diferencia del OCR tradicional, Textract identifica automáticamente la estructura del documento, detectando tablas, campos de formularios, encabezados y secciones anidadas, exportando los resultados en formato estructurado (JSON).

Capacidades técnicas:

- Detección de relaciones campo/valor en formularios complejos.
- Extracción simultánea de texto y geometría para mapeo visual.
- Procesamiento masivo mediante integración con **Amazon S3** y **AWS Lambda** para activación automática.

4.9.2. Transcripción Automática de Audio

Este servicio está respaldado por **Amazon Transcribe**, el cual permite convertir grabaciones o flujos de audio en texto estructurado con alta precisión fonética y contextual. La transcripción es multilenguaje y permite personalización mediante vocabularios específicos.

Características técnicas:

- Transcripción en lote o en tiempo real (streaming).
- Detección automática de locutores (speaker diarization).
- Procesamiento por evento mediante integración con **Amazon Kinesis** y activación por **AWS Lambda**.

4.9.3. Procesamiento de Lenguaje Natural (NLP)

El procesamiento semántico y lingüístico está basado en **Amazon Comprehend**, que ofrece análisis de sentimientos, detección de entidades, clasificación de temas, y

© Blend 2026. Todos los derechos reservados.

segmentación de textos en frases clave. También se emplean embeddings vectoriales para tareas avanzadas como detección de similitud semántica.

Servicios involucrados:

- **Comprehend**: extracción contextual de entidades y relaciones.
- **Comprehend Custom**: entrenamiento con dominios específicos (jurídico, médico, regulatorio).
- **SageMaker**: para despliegue de modelos NLP personalizados o fine-tuned.

4.9.4. Detección de Entidades y Patrones

Este servicio combina los resultados de Comprehend con modelos entrenados en **Amazon SageMaker**, permitiendo el reconocimiento de patrones complejos, como correlaciones entre entidades o estructura lógica de textos narrativos. También se integra con **OpenSearch** (antes Elasticsearch) para exploración semántica y agrupación por tópicos.

Capacidades adicionales:

- Generación de vectores semánticos.
- Indexación por similitud contextual usando **k-NN Search en OpenSearch**.
- Visualización de entidades en grafos mediante componentes personalizados.

4.9.5. Pipelines Multimedia Automatizados

Estos flujos utilizan **AWS Step Functions** para orquestar tareas complejas sobre documentos mixtos. Combinan llamadas a:

- **Amazon Textract** (extracción de texto e imagen),
- **Amazon Transcribe** (audio a texto),
- **Amazon Rekognition** (análisis visual),
- y **AWS Glue** (estructuración y transformación de salida).

La entrada y salida de datos se gestiona sobre **Amazon S3**, y los resultados se almacenan o indexan para consulta posterior, todo bajo esquemas definidos de partición y versionamiento.

4.9.6. Migración de Datos y Documentos

Este servicio (ampliado en las capacidades de migración, HM-6) está basado en **AWS DataSync**, que permite mover grandes volúmenes de datos desde sistemas locales (on-premise) hacia Amazon S3 o Amazon FSx, con alta eficiencia y bajo impacto en sistemas de origen. Se utiliza también **AWS Database Migration Service (DMS)** para sincronización o transferencia de datos estructurados desde bases relacionales.

Características técnicas:

- Validación de integridad y cifrado en tránsito.
- Soporte para planificación, reintentos y recuperación ante fallos.
- Registro de trazabilidad mediante **Amazon CloudWatch Logs**.

© Blend 2026. Todos los derechos reservados.

Cada uno de estos servicios está diseñado para operar bajo esquemas **serverless** o **gestionados**, lo que elimina la necesidad de mantener infraestructura dedicada, garantiza escalabilidad automática y permite ejecutar flujos de alto volumen sin cuellos de botella.

4.9.7. Personalización e Integración Extendida

Houndoc ha sido desarrollado con una arquitectura modular, flexible y desacoplada, que permite adaptarse dinámicamente a las necesidades particulares de cada organización. Esta flexibilidad le permite actuar como una plataforma evolutiva que puede ser **personalizada e integrada con cualquier servicio de AWS**, según los requerimientos funcionales, operativos o normativos del entorno en el que se despliegue.

Gracias a su diseño basado en componentes interoperables y flujos orquestados, Houndoc puede incorporar nuevos módulos o capacidades sin modificar su núcleo, manteniendo la estabilidad de la solución mientras se amplía su alcance. Esto permite, por ejemplo:

- Extender la plataforma con servicios de inteligencia artificial, automatización, integración de datos, monitoreo o seguridad.
- Adaptar los flujos de procesamiento y análisis a dominios específicos mediante lógica configurable.
- Desplegar funcionalidades adicionales conforme evolucionan las necesidades del negocio o los marcos regulatorios.

Asimismo, su compatibilidad con esquemas multiambiente (cloud, híbrido, on-premise) y su capacidad de integración mediante APIs, eventos y conectores, hacen posible la incorporación de **cualquier servicio de AWS**, sin restricciones de arquitectura o tecnología, permitiendo a Houndoc escalar en complejidad funcional con rapidez y eficiencia.

Esta capacidad de personalización garantiza que la plataforma no solo responda a las necesidades actuales, sino que también esté preparada para evolucionar junto con la estrategia tecnológica de la organización.

4.10. Beneficios en datos, analítica e IA

Houndoc es una plataforma integral que actúa como **acelerador para la implementación de proyectos de inteligencia artificial y analítica avanzada**, permitiendo a las organizaciones reducir drásticamente los tiempos de desarrollo, despliegue y adopción de soluciones inteligentes. Su diseño modular, su arquitectura basada en microservicios, y su profunda integración con servicios nativos de AWS la convierten en una herramienta clave para transformar rápidamente procesos intensivos en datos en sistemas automatizados, escalables y orientados a resultados.

A continuación, se detallan los principales beneficios que aporta la solución, tanto desde el punto de vista tecnológico como operativo y estratégico:

© Blend 2026. Todos los derechos reservados.

Eficiencia Operativa y Automatización Inteligente

Houndoc optimiza los flujos de procesamiento documental, extracción de datos, clasificación, validación y análisis, reduciendo la intervención humana en tareas repetitivas o de bajo valor. Mediante flujos orquestados, modelos pre-entrenados y procesamiento cognitivo automatizado, la plataforma permite acelerar operaciones que antes tomaban horas o días, llevándolas a segundos o minutos, con resultados más confiables y consistentes.

Aceleración de Proyectos de IA y Analítica

Gracias a su arquitectura preconstruida, basada en componentes reutilizables e integrados con servicios avanzados de AWS como Amazon Bedrock, SageMaker, Textract, Comprehend y Glue, Houndoc permite desplegar proyectos de inteligencia artificial, automatización documental y análisis avanzado en tiempos significativamente reducidos. Lo que normalmente requeriría meses de desarrollo, puede configurarse y poner en marcha en semanas o incluso días, sin comprometer calidad ni seguridad.

Escalabilidad Transparente y Arquitectura Serverless

La plataforma aprovecha modelos de ejecución sin servidor (serverless) y servicios gestionados, lo que permite escalar automáticamente en función de la demanda sin intervención operativa. Componentes como AWS Lambda, Glue, Step Functions, Kinesis y S3 garantizan elasticidad en la ejecución de flujos, resiliencia ante cargas variables y optimización de costos, permitiendo operar desde pequeños volúmenes hasta grandes despliegues institucionales.

Analítica de Alta Precisión y Modelado Predictivo

Houndoc permite transformar datos estructurados y no estructurados en conocimiento accionable mediante técnicas de procesamiento de lenguaje natural, visión computacional, análisis semántico y machine learning. Sus modelos pueden ser personalizados por dominio, adaptados mediante fine-tuning, o integrados con flujos de RAG para brindar respuestas precisas y contextualizadas, mejorando la calidad de las decisiones y reduciendo la incertidumbre operativa.

Seguridad Avanzada y Cumplimiento Normativo

Toda la plataforma está construida bajo los principios de **Security by Design**, incluyendo cifrado en tránsito y en reposo, autenticación federada (OpenID, OAuth 2.0), gestión de acceso basada en roles (RBAC) y auditoría centralizada con CloudTrail y CloudWatch. Houndoc puede operar bajo marcos normativos exigentes como ISO 27001, HIPAA, GDPR o regulaciones nacionales de protección de datos, lo que la hace ideal para sectores como salud, justicia, gobierno o servicios financieros.

Reducción de Errores y Mejora en la Confiabilidad

La automatización de procesos críticos reduce errores humanos, estandariza resultados y permite monitorear la ejecución de cada flujo en tiempo real. Las validaciones automáticas, el control de calidad de entradas, y el versionamiento de flujos y modelos aseguran consistencia, repetibilidad y trazabilidad total.

© Blend 2026. Todos los derechos reservados.

Integración Abierta y Sin Fricciones

Houndoc está diseñado como un entorno interoperable, que se puede integrar con bases de datos existentes, herramientas BI, motores de búsqueda, portales de usuario o plataformas transaccionales. Gracias al uso de APIs RESTful, conectores nativos y eventos, puede operar como parte de ecosistemas tecnológicos ya establecidos, sin necesidad de reemplazo o reestructuración total.

Flexibilidad Modular y Escenarios Personalizables

La plataforma está compuesta por módulos independientes y parametrizables, lo que permite activar únicamente las funcionalidades requeridas para cada caso de uso. Esto reduce complejidad inicial, permite adopción progresiva, y facilita la adaptación a múltiples industrias, regulaciones, idiomas y modelos operativos.

Alineación Estratégica y Retorno Acelerado

Al integrar en una sola plataforma todas las piezas necesarias para digitalizar, automatizar y analizar información crítica, Houndoc actúa como catalizador de la transformación digital, alineando capacidades técnicas con los objetivos estratégicos de cada organización. Su enfoque orientado a resultados permite obtener beneficios tangibles desde las primeras etapas del proyecto, con mejoras progresivas en eficiencia, calidad, cumplimiento y toma de decisiones basadas en datos.

4.11. Transformación digital acelerada

Houndoc representa una solución estratégica para organizaciones que buscan acelerar la implementación de inteligencia artificial, automatización y analítica avanzada con una infraestructura confiable, segura y escalable. Su diseño modular, sustentado por servicios nativos de AWS, permite abordar de forma integral los desafíos asociados al procesamiento de información crítica, la gestión documental compleja y la transformación de datos en conocimiento accionable.

Al combinar componentes funcionales altamente especializados con capacidades cognitivas de última generación, Houndoc se posiciona como un habilitador real de la transformación digital institucional. Ofrece no solo tecnología avanzada, sino también una arquitectura flexible que se adapta a distintos dominios operativos, regulaciones locales y niveles de madurez tecnológica.

La plataforma ha sido concebida como un **acelerador de soluciones inteligentes**, permitiendo a las organizaciones reducir barreras técnicas, optimizar tiempos de desarrollo y asegurar la sostenibilidad operativa de sus iniciativas de datos e inteligencia artificial. En contextos donde la eficiencia, la trazabilidad, la gobernanza de datos y la toma de decisiones basada en evidencia son imperativos estratégicos, Houndoc entrega una respuesta integral, robusta y comprobada.

Su enfoque escalable y seguro permite iniciar con proyectos específicos de alto impacto, con capacidad de evolucionar hacia despliegues institucionales más amplios, apoyando la visión de construir organizaciones más inteligentes, ágiles y centradas en el conocimiento.

© Blend 2026. Todos los derechos reservados.

5. Capacidades de migración y modernización

5.1. Houndoc como acelerador de migraciones

Houndoc es el acelerador de inteligencia artificial de Blend para migrar y modernizar cargas de trabajo hacia la nube de AWS. Permite ejecutar la migración dentro de plazos exigentes **sin modificar las aplicaciones migradas**. Automatiza las tareas de análisis, diseño, planeación y validación de la migración, y reduce el tiempo que estas actividades requieren cuando se realizan de forma manual.

Houndoc combina agentes de inteligencia artificial especializados en cada fase del proyecto con componentes reutilizables y preconfigurados. Los agentes procesan inventarios, métricas y configuraciones; proponen el diseño, el dimensionamiento y el orden de la migración; orquestan y validan cada ola, y documentan el resultado. Los equipos técnicos revisan y aprueban cada decisión antes de que se aplique.

Cuando la organización decide ir más allá del traslado, Houndoc también acelera la **modernización de aplicaciones**: análisis y documentación de código legado, actualización de frameworks, conversión de bases de datos y transformación de aplicaciones hacia arquitecturas nativas de la nube.

5.2. El reto de migrar y modernizar

Las migraciones a la nube se retrasan o fallan por causas que se repiten de un proyecto a otro:

- **Inventarios incompletos o desactualizados.** La información de servidores, almacenamiento y aplicaciones está dispersa entre consolas de virtualización, CMDB y hojas de cálculo.
- **Dependencias ocultas.** Las comunicaciones entre aplicaciones, servidores y sistemas externos no están documentadas, y se descubren cuando una migración rompe una conexión crítica.
- **Dimensionamiento a ojo.** Se replica en la nube la capacidad asignada en el centro de datos, no la que realmente se consume.
- **Planeación manual.** Las olas, los planes de contingencia y los runbooks de corte se construyen a mano para cada proyecto.
- **Validación lenta.** Las pruebas de conectividad, configuración e integridad de datos se ejecutan una por una y se convierten en el cuello de botella de cada ventana de mantenimiento.
- **Documentación al final.** La memoria técnica y la transferencia de conocimiento se elaboran al cierre, con información incompleta.
- **Código legado.** Aplicaciones en .NET Framework, mainframe o versiones antiguas de Java, y bases de datos comerciales que no aprovechan la nube sin una modernización costosa y riesgosa.

© Blend 2026. Todos los derechos reservados.

5.3. Cómo funciona: agentes de IA con aprobación humana

Houndoc organiza el proyecto como un conjunto de trabajos ejecutados por agentes de IA especializados: descubrimiento y evaluación, red y Landing Zone, planeación, migración de servidores, validación, bases de datos, modernización y documentación. Cada agente forma parte de un módulo de migración (sección 5.8) y:

- Recibe la información de entrada: inventarios, exportaciones de herramientas, configuraciones o código fuente.
- Genera una propuesta (diseño, plan, código o documentación) y explica su razonamiento.
- Se detiene en **puntos de revisión** donde el equipo técnico del cliente y de Blend aprueba, ajusta o rechaza la propuesta.
- Ejecuta solo lo aprobado y registra cada acción para su auditoría.

El equipo interactúa con los agentes desde un espacio de trabajo por proyecto, mediante una interfaz conversacional en lenguaje natural: puede pedir explicaciones, cambiar prioridades, excluir cargas o solicitar alternativas sin reprogramar el flujo.

5.4. Capacidades (HM)

Las capacidades se agrupan en ocho familias. Cada una indica el módulo que la implementa.

HM-1. Descubrimiento automatizado

Módulo: MDA - Migration Discovery & Assessment

- **HM-1.1. Inventario de activos.** Procesa exportaciones de VMware vCenter/vSAN, Microsoft Hyper-V, servidores físicos, consolas de almacenamiento, CMDB, herramientas de descubrimiento y hojas de cálculo. Normaliza la información en un inventario único de servidores, máquinas virtuales, discos, sistemas de almacenamiento de bloque, NAS y datos no estructurados, con su ubicación, y señala los datos faltantes o inconsistentes.
- **HM-1.2. Mapa de dependencias.** Con los datos de conexiones de red, las configuraciones y la información de aplicaciones proporcionada por el cliente, genera el mapa de comunicación entre servidores, aplicaciones y sistemas externos. Identifica las cargas que deben migrarse en conjunto, las dependencias con otras entidades y las aplicaciones de uso interno que no requieren publicarse a internet.
- **HM-1.3. Dimensionamiento basado en consumo real.** Analiza las métricas de uso de CPU, memoria, almacenamiento, IOPS y red, y recomienda para cada carga el tipo de instancia de cómputo, el tipo de disco y el nivel de almacenamiento, en lugar de replicar la capacidad asignada en origen.
- **HM-1.4. Evaluación de preparación.** Determina qué cargas se pueden migrar directamente y cuáles requieren cambios, clasifica cada una por estrategia de migración (rehost, replatform, refactor, retire, retain) y detecta bloqueadores técnicos como sistemas operativos fuera de soporte o licenciamiento restrictivo.

© Blend 2026. Todos los derechos reservados.

- **HM-1.5. Análisis de almacenamiento y ancho de banda.** Clasifica el almacenamiento por rendimiento (IOPS para bloque), capacidad (NAS) y criticidad de acceso (datos no estructurados), y estima el ancho de banda, la latencia y los tiempos de transferencia necesarios para cumplir con el calendario.
- **HM-1.6. Seguridad y servicios complementarios.** A partir de las exportaciones de configuración de red y seguridad, identifica firewalls, reglas de acceso, balanceadores de carga, direccionamiento, zonas de red, publicación de aplicaciones, resolución de nombres y servicios de directorio que deben considerarse en el diseño.
- **HM-1.7. Caso de negocio.** Estima el costo de operación en la nube bajo distintas opciones de compra (bajo demanda, Savings Plans e instancias reservadas) y alternativas de licenciamiento de Windows Server y SQL Server, como insumo para la decisión de migrar.

HM-2. Diseño de arquitectura como código

Módulo: NLZ - Network & Landing Zone

- **HM-2.1. Traducción de la red de origen.** Interpreta la configuración de red de origen, como exportaciones de VMware NSX y vSphere (segmentos, grupos, reglas de firewall, NAT y balanceadores), y genera su equivalente en AWS: VPC, subredes, tablas de ruteo, grupos de seguridad, listas de control de acceso y NAT gateways, conservando la segmentación y las reglas de comunicación vigentes.
- **HM-2.2. Landing Zone como código.** Genera la arquitectura de destino a partir de plantillas reutilizables que incluyen estructura de cuentas, zonas de red, identidades, políticas de seguridad y etiquetado de recursos: estructura multicuenta con AWS Organizations, políticas de control de servicio, identidades con AWS IAM Identity Center y registro centralizado con AWS CloudTrail y AWS Config. El resultado se entrega como infraestructura como código en AWS CloudFormation, AWS CDK o Terraform, sujeto a revisión y aprobación antes de su despliegue.
- **HM-2.3. Estrategias de servicio.** Propone el mapeo de cada máquina virtual a su equivalente en la nube; del almacenamiento de bloque a discos administrados; de NAS a servicios de archivos administrados, y de datos no estructurados a almacenamiento de objetos con políticas de ciclo de vida (capas caliente, fría y de archivo). Incluye en el código de la Landing Zone las estrategias de respaldo, monitoreo, seguridad y automatización.
- **HM-2.4. Conectividad, direccionamiento y dependencias externas.** Diseña la conectividad híbrida desde los centros de datos hacia la nube con AWS Transit Gateway, enlaces dedicados o VPN; el direccionamiento para aplicaciones internas y publicadas a internet, y la conectividad necesaria para mantener las dependencias con entidades internas y externas.
- **HM-2.5. Validación del diseño.** Revisa el diseño contra reglas de seguridad, disponibilidad y costo, y reporta las desviaciones antes de su presentación para aprobación.

© Blend 2026. Todos los derechos reservados.

HM-3. Planeación de olas

Módulo: WPL - Wave Planning

- **HM-3.1. Estrategia de olas.** Agrupa servidores y almacenamiento en bloques de migración según su complejidad, dependencias, criticidad y ventanas de mantenimiento autorizadas, comenzando por entornos de prueba o aplicaciones no críticas para estabilizar la plataforma antes de llegar a las aplicaciones críticas.
- **HM-3.2. Plan de contingencia y retorno.** Genera para cada bloque su plan de contingencia y su estrategia de retorno en caso de que la migración falle durante la ventana de mantenimiento.
- **HM-3.3. Runbooks de corte.** Produce el runbook de cada ventana: secuencia de actividades, responsables, criterios de avance y de retorno, cambios de ruteo y verificaciones posteriores.
- **HM-3.4. Programa de trabajo.** Consolida las olas en un programa con ruta crítica, actividades del proveedor y del cliente, y fechas de ejecución.
- **HM-3.5. Selección de herramientas de replicación.** Recomienda el método de replicación para cada tipo de carga: replicación continua a nivel de bloque para servidores, transferencia de archivos para NAS y datos no estructurados, replicación de bases de datos, y dispositivos físicos cuando el volumen excede la capacidad de los enlaces.

HM-4. Ejecución y replicación

Módulos: SMO - Server Migration Orchestration, NLZ, DBM y Data Ingestion Management

- **HM-4.1. Creación de la Landing Zone.** Despliega la Landing Zone aprobada a partir de sus archivos de infraestructura como código, de forma controlada y repetible.
- **HM-4.2. Permisos de las cuentas de servicio.** Opera con roles de permisos acotados y genera el reporte de los permisos otorgados al motor de ejecución de la infraestructura como código, para auditorías de seguridad.
- **HM-4.3. Instalación de agentes y configuración de la replicación.** Configura la replicación de servidores con agente en el servidor de origen o sin agente para entornos VMware vCenter, sobre la conectividad establecida hacia la nube.
- **HM-4.4. Replicación inicial.** Orquesta la copia del grueso de los datos (máquinas virtuales, servidores físicos, archivos de NAS, almacenamiento de bloque y bases de datos) hacia la nube en segundo plano.
- **HM-4.5. Sincronización delta.** Mantiene los datos en la nube actualizados con los cambios que ocurren en los servidores y bases de datos de origen hasta el momento del corte.
- **HM-4.6. Migración de entornos VMware.** Acelera la salida de entornos VMware: traduce su red, agrupa las máquinas virtuales por dependencias y orquesta su migración por olas.

© Blend 2026. Todos los derechos reservados.

HM-5. Validación automatizada

Módulo: MVA - Migration Validation

- **HM-5.1. Pruebas de aceptación en entornos aislados.** Lanza instancias de prueba con los datos replicados y ejecuta pruebas de conectividad, puertos, servicios, configuración del sistema operativo, almacenamiento, monitoreo, respaldos, balanceo y seguridad.
- **HM-5.2. Validación de la infraestructura en zonas de disponibilidad.** Verifica mediante comandos la configuración de la infraestructura desplegada en cada zona de disponibilidad.
- **HM-5.3. Integridad de datos.** Verifica la integridad de los archivos y bases de datos migrados, con los reportes de verificación de la transferencia y la comparación de datos entre origen y destino.
- **HM-5.4. Validación en paralelo y evidencia.** Valida varias olas al mismo tiempo, con flujos desacoplados, y genera para cada una un paquete de evidencia para la aceptación del cliente.
- **HM-5.5. Soporte a la ventana de transición.** Ejecuta la sincronización final, el lanzamiento de corte y las verificaciones posteriores al cambio de ruteo (DNS/IP) definidas en el runbook.

HM-6. Migración de datos y contenido

Módulos: DBM - Database Migration, Data Ingestion Management, Data Lake, ICG y AAN

- **HM-6.1. Transferencia de archivos y almacenamiento.** Migra grandes volúmenes desde NAS, servidores de archivos y almacenamiento local hacia almacenamiento de objetos y servicios de archivos administrados, con programación, control de ancho de banda, validación de integridad y cifrado en tránsito.
- **HM-6.2. Migración y replicación de bases de datos.** Ejecuta cargas iniciales y replicación continua desde motores relacionales, sin afectar la disponibilidad del origen y con validación de los datos migrados.
- **HM-6.3. Conversión de esquemas entre motores.** Convierte esquemas, vistas, procedimientos almacenados y funciones de bases de datos comerciales hacia motores de código abierto o administrados, con apoyo de IA generativa, y reporta los elementos que requieren intervención manual.
- **HM-6.4. Catalogación, clasificación y linaje.** Cataloga los activos de datos, detecta información sensible y registra el linaje desde el origen hasta el destino.
- **HM-6.5. Data lake de destino.** En migraciones de plataformas analíticas, organiza los datos migrados en zonas de datos crudos, depurados y curados, con transformación automatizada y catálogo central.
- **HM-6.6. Digitalización de repositorios documentales y multimedia.** Extrae texto y estructura de documentos escaneados, transcribe audio y video, y clasifica el contenido migrado para su indexación y búsqueda.

© Blend 2026. Todos los derechos reservados.

HM-7. Documentación y transferencia

Módulos: MKD - Migration Knowledge & Documentation, MDA y Dashboards

- **HM-7.1. Documentación de cada bloque migrado.** Genera la memoria técnica as-built de cada ola: inventario final, diagramas, configuración aplicada, resultados de validación e incidencias atendidas. Sirve como base para el acta de entrega-recepción.
- **HM-7.2. Materiales de transferencia de conocimiento.** Produce temarios, guías y materiales de capacitación sobre la plataforma de destino para los equipos que la operarán.
- **HM-7.3. Asistente y tablero de avance.** Un asistente conversacional responde preguntas sobre el inventario, las dependencias, el plan y el estado del proyecto, y un tablero muestra el avance de cada ola contra el programa.
- **HM-7.4. Informes de consumo y optimización.** Analiza el consumo real en la nube, explica desviaciones de costo y recomienda ajustes de dimensionamiento de servidores y la depuración de almacenamiento subutilizado o sin uso.

HM-8. Modernización de aplicaciones

Módulo: MOD - Application Modernization. Es un camino opcional, separado de la migración sin cambios.

- **HM-8.1. Análisis y documentación de código legado.** Identifica componentes, dependencias, reglas de negocio y flujos de datos, y genera documentación técnica y funcional.
- **HM-8.2. Modernización de aplicaciones .NET.** Transforma aplicaciones .NET Framework hacia .NET multiplataforma para Linux, con compilación y ejecución de pruebas unitarias.
- **HM-8.3. Modernización de mainframe.** Analiza COBOL, JCL, CICS, Db2 y VSAM; extrae reglas de negocio; descompone el sistema en dominios; planifica olas de modernización; refactoriza hacia Java, y genera planes de prueba y scripts de captura de datos de prueba.
- **HM-8.4. Modernización de aplicaciones Windows de pila completa.** Moderniza en conjunto la aplicación, su base de datos SQL Server y su despliegue, hacia .NET multiplataforma, Amazon Aurora PostgreSQL y contenedores o Linux.
- **HM-8.5. Actualización de versiones y frameworks.** Actualiza aplicaciones Java, Python y Node.js a versiones con soporte vigente, junto con sus librerías y SDK.
- **HM-8.6. Transformaciones personalizadas.** Aplica reglas de transformación definidas en lenguaje natural o con ejemplos de forma consistente a múltiples repositorios, y mejora con la retroalimentación del equipo.
- **HM-8.7. Contenerización.** Empaqueta aplicaciones para Amazon ECS o Amazon EKS y genera sus artefactos de despliegue.
- **HM-8.8. Pruebas y equivalencia.** Genera pruebas para el código transformado y verifica que su comportamiento sea equivalente al original. Cada

© Blend 2026. Todos los derechos reservados.

transformación se entrega en una rama nueva, con resumen de cambios y resultados de compilación y pruebas.

5.5. Fundamentos de aceleración (HF)

ID	Fundamento	Cómo lo cumple Houndoc	Módulos
HF-1	Componentes reutilizables y preconfigurados: descubrimiento, procesamiento, validación y documentación no se desarrollan desde cero para cada proyecto	Agentes, conectores, plantillas de Landing Zone, flujos y formatos de documentación preconstruidos	MDA, NLZ, MVA, MKD, WFM
HF-2	Automatización de las tareas que de forma manual toman horas o días: extracción de datos, clasificación, validación y análisis de dependencias	Agentes de IA que procesan inventarios, dependencias y validaciones con resultados consistentes, liberando al equipo para las decisiones que requieren juicio humano	MDA, MVA, AIG
HF-3	Arquitectura modular desplegada sobre servicios gestionados sin servidor, con capacidad ajustable al volumen sin intervención manual	AWS Step Functions, AWS Lambda, Amazon EventBridge y Amazon Bedrock escalan automáticamente	WFM, AIG
HF-4	Ejecución en lotes y en paralelo: varios bloques se validan al mismo tiempo	Ejecución simultánea de olas de replicación y de validación	WFM, SMO, MVA
HF-5	Flujos de trabajo desacoplados: una falla en un bloque no detiene el resto	Cada flujo y cada paso se ejecuta de forma aislada, con reintentos y manejo de errores	WFM
HF-6	Cada flujo es una especificación versionable y auditable, monitoreada de forma independiente	Flujos declarativos con control de versiones; métricas, registros y alarmas por flujo en Amazon CloudWatch	WFM
HF-7	Seguridad desde el diseño: cifrado en tránsito y en reposo,	Llaves administradas con AWS KMS, federación OpenID	IAM, MWS

© Blend 2026. Todos los derechos reservados.

ID	Fundamento	Cómo lo cumple Houndoc	Módulos
	autenticación federada, control de acceso por roles y auditoría centralizada	Connect / OAuth 2.0 / SAML 2.0, roles por proyecto, AWS CloudTrail	
HF-8	Validaciones automáticas de entrada, control de calidad de los datos y versionamiento de flujos y modelos para reducir errores humanos y mantener la trazabilidad	Validación de entradas en la ingesta, postprocesamiento automático de los resultados de los agentes, catálogo y linaje, bitácora de trabajo	Data Ingestion, AIG, ICG, MWS
HF-9	Migración de datos con validación de integridad, reintentos ante fallas y registro de cada transferencia	Verificación de transferencias de archivos, validación de datos de bases de datos, reintentos automáticos y registro en CloudWatch Logs	Data Ingestion, DBM, MVA
HF-10	Módulos independientes y parametrizables: se activan solo las capacidades necesarias para cada caso	Cada módulo se habilita por proyecto	Todos
HF-11	Despliegue en nube en la región de México	Disponible en la región de AWS México (Central), además de Norte de Virginia y otras regiones	Todos

5.6. Houndoc en cada fase del proyecto de migración

La siguiente tabla relaciona las fases y actividades típicas de un proyecto de migración con las capacidades y módulos de Houndoc que las soportan.

Fase	Actividad	Capacidades	Módulos
1. Descubrimiento de la infraestructura a migrar	Inventario de activos	HM-1.1	MDA
	Dependencias de aplicaciones, internas y externas	HM-1.2	MDA

© Blend 2026. Todos los derechos reservados.

Fase	Actividad	Capacidades	Módulos
	Evaluación de preparación	HM-1.4	MDA
	Análisis de almacenamiento y anchos de banda	HM-1.5	MDA
	Seguridad y servicios complementarios	HM-1.6	MDA, NLZ
2. Diseño de la arquitectura en la nube	Landing Zone: cuentas, redes, subredes, políticas de seguridad y gobernanza	HM-2.1, HM-2.2	NLZ
	Dimensionamiento	HM-1.3	MDA
	Estrategias de virtualización, almacenamiento (bloque, NAS, no estructurado), respaldos, monitoreo, seguridad y automatización	HM-2.3, HM-2.2	NLZ
	Dependencias externas y conectividad	HM-2.4	NLZ
	Revisión y aprobación del diseño	HM-2.5	NLZ, MWS
3. Planificación de la migración	Estrategia de olas	HM-3.1	WPL
	Selección de herramientas	HM-3.5	WPL
	Plan de contingencia y retorno	HM-3.2	WPL
	Direccionamiento para la publicación de aplicaciones	HM-2.4	NLZ
	Programa de trabajo	HM-3.4	WPL
4. Ejecución y replicación previa	Creación de la Landing Zone con herramienta de automatización	HM-4.1	NLZ
	Permisos de las cuentas de servicio del motor de infraestructura como código	HM-4.2	NLZ, IAM
	Instalación de agentes y configuración de red	HM-4.3	SMO

© Blend 2026. Todos los derechos reservados.

Fase	Actividad	Capacidades	Módulos
	Replicación inicial	HM-4.4, HM-6.1, HM-6.2	SMO, DBM, Data Ingestion
	Sincronización delta	HM-4.5	SMO, DBM, Data Ingestion
	Validación de la infraestructura en zonas de disponibilidad	HM-5.2	MVA
	Memoria técnica de implementación	HM-7.1	MKD
5. Pruebas y transición	Pruebas de aceptación en entornos aislados	HM-5.1	SMO, MVA
	Integridad de los datos	HM-5.3	MVA
	Ventana de transición	HM-3.3, HM-5.5	WPL, SMO, MVA
	Validación post-migración y aceptación por bloque	HM-5.4	MVA
6. Migración y entrega	Documentación de cada bloque y acta de entrega-recepción	HM-7.1	MKD
	Aseguramiento y gobierno: respaldos, monitoreo y seguridad	HM-2.3	NLZ
	Transferencia de conocimiento	HM-7.2	MKD
	Informes de consumo y recomendaciones de ahorro	HM-7.4	MDA, MKD
Modernización (opcional)	Análisis, transformación y pruebas de aplicaciones y bases de datos	HM-8.1 a HM-8.8, HM-6.3	MOD, DBM

5.7. Cargas, plataformas y servicios que cubre

Dominio	Alcance
Cómputo	Máquinas virtuales VMware vSphere/vSAN y Microsoft Hyper-V; servidores físicos, incluidos servidores de alta capacidad; sistemas operativos Windows Server y Linux
Almacenamiento	Almacenamiento de bloque (SAN, vSAN); NAS y servidores de archivos; datos no

© Blend 2026. Todos los derechos reservados.

Dominio	Alcance
	estructurados, incluidos repositorios documentales, video y audio
Bases de datos	Motores relacionales comerciales y de código abierto, con migración homogénea o conversión hacia otros motores
Red y conectividad	Segmentos y direccionamiento, reglas de firewall, balanceo de cargas, resolución de nombres, conectividad híbrida y dependencias internas y externas
Seguridad e identidades	Estructura de cuentas, identidades y roles, políticas de seguridad, cifrado y etiquetado de recursos
Aplicaciones	Migración sin cambios de cualquier aplicación sobre las plataformas anteriores; modernización de .NET, mainframe, Java y aplicaciones Windows de pila completa

Houndoc acelera el trabajo de migración y modernización. No sustituye las herramientas de replicación seleccionadas para el proyecto ni los servicios de infraestructura de nube que el cliente contrata para operar sus cargas, ni las mesas de trabajo, validaciones y aprobaciones que el cliente establece para cada fase.

5.8. Módulos de migración y modernización

MWS - Migration Workspace Es el espacio de trabajo colaborativo de cada proyecto de migración. Agrupa los trabajos (jobs) del proyecto, como una evaluación, una ola de migración o la modernización de un repositorio, y permite al equipo interactuar con los agentes mediante chat en lenguaje natural. Administra los puntos de revisión donde los usuarios autorizados aprueban, ajustan o rechazan cada propuesta, y lleva una bitácora de trabajo con cada acción de los agentes y de las personas. Asigna roles por proyecto (administración, colaboración, aprobación y consulta) y gestiona los conectores hacia las cuentas de AWS de destino, mediante roles con permisos acotados, y hacia los repositorios de código. Los artefactos generados se guardan cifrados con llaves administradas por AWS Key Management Service.

MDA - Migration Discovery & Assessment Es el módulo de descubrimiento y evaluación. Importa exportaciones de inventario de VMware vCenter, Microsoft Hyper-V, herramientas de descubrimiento, CMDB y hojas de cálculo, y las normaliza en un inventario único de servidores, discos, redes y aplicaciones, con reporte de datos faltantes o inconsistentes. Con los datos de conexiones de red construye el grafo de dependencias entre servidores, aplicaciones y sistemas externos, y agrupa las cargas que deben migrarse juntas. A partir de las métricas de utilización recomienda el tipo de instancia de Amazon EC2 y de volumen de Amazon EBS para cada servidor, clasifica cada carga por estrategia de migración y detecta bloqueadores técnicos. Genera el caso de negocio con el costo estimado en la nube bajo distintas opciones de compra (bajo demanda, Savings Plans e instancias reservadas) y alternativas de licenciamiento de Windows Server y SQL Server. Después de la migración, repite la evaluación con las

© Blend 2026. Todos los derechos reservados.

métricas reales de la nube para recomendar ajustes de dimensionamiento y detectar almacenamiento subutilizado.

NLZ - Network & Landing Zone Es el módulo de diseño de la arquitectura de destino. Interpreta la configuración de red de origen, como exportaciones de VMware NSX y vSphere (segmentos, grupos, reglas de firewall, NAT y balanceadores), y genera la red equivalente en AWS: VPC, subredes, tablas de ruteo, grupos de seguridad, listas de control de acceso, NAT gateways y AWS Transit Gateway para la conectividad entre redes y con los centros de datos. Construye la Landing Zone a partir de plantillas reutilizables: estructura multicuenta con AWS Organizations y unidades organizacionales, políticas de control de servicio, identidades con AWS IAM Identity Center, registro centralizado con AWS CloudTrail y AWS Config, y políticas de etiquetado. Entrega el resultado como infraestructura como código en AWS CloudFormation, AWS CDK o Terraform, lo valida contra reglas de seguridad, disponibilidad y costo, y lo despliega solo después de su aprobación.

WPL - Wave Planning Es el módulo de planeación. Usa el grafo de dependencias, la complejidad, la criticidad y las ventanas de mantenimiento autorizadas para agrupar las cargas en olas y ordenarlas, comenzando por entornos de prueba o aplicaciones no críticas. Permite ajustar tamaño, orden y contenido de cada ola mediante conversación con el agente. Para cada ola genera el plan de contingencia, la estrategia de retorno y el runbook de corte, y consolida todo en un programa de trabajo con ruta crítica y responsables. Las olas aprobadas se sincronizan con el módulo SMO para su ejecución.

SMO - Server Migration Orchestration Es el módulo de migración de servidores. Orquesta AWS Application Migration Service para replicar servidores físicos y máquinas virtuales Windows y Linux mediante replicación continua a nivel de bloque, con agente en el servidor de origen o sin agente para entornos VMware vCenter. Configura las plantillas de lanzamiento de cada servidor con el dimensionamiento de MDA y la red de NLZ, lanza instancias de prueba en entornos aislados, ejecuta el corte de cada ola, aplica acciones automatizadas posteriores al lanzamiento con AWS Systems Manager y cierra la replicación de los servidores ya migrados. Ejecuta varias olas en paralelo y reporta su estado en tiempo real.

MVA - Migration Validation Es el módulo de validación. Después de cada lanzamiento de prueba o de corte, ejecuta con AWS Systems Manager pruebas de conectividad y puertos, estado de servicios, configuración del sistema operativo y verificación de archivos sobre las instancias migradas, y comprueba la configuración de la infraestructura en las zonas de disponibilidad. Incorpora los reportes de verificación de AWS DataSync y la validación de datos de AWS Database Migration Service. Ejecuta las validaciones de varias olas en paralelo y genera para cada una un paquete de evidencia para la aceptación del cliente.

DBM - Database Migration Es el módulo de migración de bases de datos. Evalúa las bases de datos de origen y su complejidad de migración. Para migraciones entre motores distintos, como Oracle o SQL Server hacia Amazon Aurora PostgreSQL, Amazon RDS o MySQL, convierte esquemas, vistas, procedimientos almacenados y funciones con apoyo de IA generativa, y entrega un reporte de los elementos que requieren intervención manual. Ejecuta la carga inicial y la replicación continua de

© Blend 2026. Todos los derechos reservados.

cambios con AWS Database Migration Service, sin afectar la disponibilidad del origen, y valida que los datos de origen y destino coincidan.

MOD - Application Modernization Es el módulo de modernización de aplicaciones. Se conecta a repositorios de código (GitHub, GitLab, Bitbucket y otros), analiza dependencias y componentes, y ejecuta transformaciones con revisión humana:

- Aplicaciones .NET Framework hacia .NET multiplataforma para Linux, con compilación y ejecución de pruebas unitarias.
- Aplicaciones Windows de pila completa: aplicación, base de datos SQL Server y despliegue, hacia .NET multiplataforma, Amazon Aurora PostgreSQL y contenedores o Linux.
- Mainframe: análisis de COBOL, JCL, CICS, Db2 y VSAM; generación de documentación técnica y funcional; extracción de reglas de negocio; descomposición en dominios; planeación de olas de modernización; refactorización hacia Java, y generación de planes de prueba y de scripts de captura de datos de prueba.
- Actualización de versiones de Java, Python y Node.js, y de sus librerías y SDK.
- Transformaciones personalizadas definidas en lenguaje natural o con ejemplos, aplicadas de forma consistente a muchos repositorios, que mejoran con la retroalimentación del equipo.
- Contenerización para Amazon ECS o Amazon EKS.

Cada transformación entrega el código en una rama nueva, con un resumen de cambios, resultados de compilación y pruebas, y los elementos pendientes de revisión.

MKD - Migration Knowledge & Documentation Es el módulo de conocimiento y documentación. Genera la memoria técnica as-built de cada ola (inventario final, diagramas, configuración aplicada, resultados de validación e incidencias), los materiales de transferencia de conocimiento y los informes de avance y de optimización. Ofrece un asistente que responde preguntas sobre el inventario, las dependencias, el plan y el estado del proyecto, mediante búsqueda semántica sobre el modelo de conocimiento del proyecto y los documentos generados.

5.9. Arquitectura de las capacidades de migración

Las capacidades de migración combinan sus módulos específicos con los módulos base de la plataforma en seis capas:

Capa	Función	Módulos
Experiencia	Espacio de trabajo del proyecto, interfaz conversacional, puntos de aprobación y tableros de avance	MWS, Dashboards
Agentes de IA	Analizan la información y generan propuestas: inventarios, diseños, olas, código y documentación	MDA, NLZ, WPL, SMO, MVA, DBM, MOD, MKD, sobre AIG

© Blend 2026. Todos los derechos reservados.

Capa	Función	Módulos
Orquestación	Ejecuta cada fase como un flujo versionable y auditable, con aprobaciones, paralelismo, reintentos y aislamiento de fallas	WFM
Ejecución e integración	Conecta con el origen y el destino: importa inventarios y código, despliega infraestructura, replica servidores, transfiere datos y ejecuta validaciones	NLZ, SMO, MVA, DBM, MOD, Data Ingestion
Conocimiento del proyecto	Modelo único con inventario, dependencias, métricas, diseños, olas, decisiones y evidencia, con su linaje	Data Lake, ICG, AAN
Seguridad y gobierno	Identidad federada, roles, cifrado y auditoría de cada acción, transversal a todas las capas	IAM

5.10. Modelo de entrega y regiones (HE)

- **HE-1.** En los proyectos de migración, Houndoc opera como herramienta de trabajo de Blend y forma parte de sus servicios de migración y modernización a la nube.
- **HE-2.** Blend despliega Houndoc en la nube, sobre su propia infraestructura, en las regiones de AWS de México (Central) y Norte de Virginia, o en otras regiones de AWS que ofrecen los servicios en los que se apoya, según lo defina el cliente.
- **HE-3.** Houndoc no consume los recursos de IaaS, PaaS y SaaS que el cliente contrata para sus cargas de trabajo ni para la migración.
- **HE-4.** Blend es responsable de la operación, el mantenimiento y la seguridad de la infraestructura de Houndoc durante la migración.
- **HE-5.** El uso de Houndoc está incluido en los servicios de migración de Blend.
- **HE-6.** Al concluir la migración, Blend entrega al cliente los resultados generados y elimina la información del cliente de la infraestructura de Houndoc, con evidencia de la eliminación.

5.11. Beneficios en migración y modernización

Plazos que se cumplen

Las tareas de análisis, diseño, planeación, validación y documentación que en una migración tradicional toman días o semanas por ola se ejecutan de forma automatizada, lo que permite cumplir calendarios exigentes sin aumentar proporcionalmente el equipo.

© Blend 2026. Todos los derechos reservados.

Menos riesgo en cada corte

El mapa de dependencias, las olas ordenadas por complejidad y criticidad, los planes de retorno y la validación automatizada reducen la probabilidad de romper conexiones críticas o perder datos durante la ventana de mantenimiento.

Costo de nube acorde al consumo real

El dimensionamiento basado en métricas y la optimización posterior evitan trasladar a la nube la sobreasignación del centro de datos.

Resultados consistentes y auditables

Cada flujo es una especificación versionada; cada decisión queda aprobada y registrada; cada transferencia tiene evidencia de integridad. El resultado es repetible entre olas y defendible ante auditorías.

Infraestructura como código desde el primer día

La Landing Zone y su evolución quedan en código revisable y reutilizable, lo que facilita la operación, el cumplimiento y los cambios futuros.

Un camino de modernización en la misma plataforma

La misma herramienta que migra las cargas sin cambios permite, cuando la organización lo decide, documentar y modernizar sus aplicaciones y bases de datos.

Continuidad hacia la explotación de los datos

Los datos migrados quedan catalogados y listos para ser explotados con las capacidades de búsqueda, analítica e IA de Houndoc (sección 4).

6. Seguridad, privacidad y gobierno

Las siguientes condiciones aplican a toda la plataforma Houndoc:

- **Seguridad desde el diseño.** Cifrado en tránsito y en reposo, autenticación federada (OpenID Connect, OAuth 2.0, SAML 2.0, LDAP), control de acceso basado en roles, auditoría completa de acciones y monitoreo en tiempo real de eventos críticos.
- **Auditoría centralizada.** Registro de accesos, cambios y operaciones con AWS CloudTrail y Amazon CloudWatch.
- **Residencia de datos.** La información del cliente se almacena y procesa en la región o el entorno seleccionado para el proyecto.
- **Los datos del cliente no se usan para entrenar modelos.** Los modelos fundacionales operan sobre servicios administrados que garantizan el aislamiento de la información.
- **Control de calidad y trazabilidad.** Validaciones automáticas de entrada, control de calidad de los datos, versionamiento de flujos y modelos, catálogo y linaje.

© Blend 2026. Todos los derechos reservados.

- **Cumplimiento normativo.** La plataforma puede operar bajo marcos como ISO 27001, HIPAA, GDPR o regulaciones nacionales de protección de datos personales.
- **Aprobación humana en migraciones.** Ninguna decisión de diseño, planeación o corte se aplica sin la aprobación del equipo responsable.

7. Propiedad intelectual

Houndoc es un programa de computación desarrollado por Blend y registrado ante el Instituto Nacional del Derecho de Autor (INDAUTOR) en México. Blend cuenta con los derechos para usarlo, reproducirlo, adaptarlo, comunicarlo y licenciarlo sin restricción de terceros en el territorio nacional, y conserva la evidencia que sustenta esos derechos: repositorio de código con historial de versiones, documentación técnica de arquitectura e inventario de componentes de terceros con sus licencias.

8. Referencias

- **Documentación técnica oficial de Houndoc:** <https://docs.houndoc.ai/>. Es la fuente de referencia para detalles de implementación, despliegue y operación de la plataforma. Secciones: Intro, User Manual, Installation Guide, Architecture Manual, Configuration Guide y Migration / Update Manual.
- **Amazon OpenSearch Service:** <https://docs.aws.amazon.com/opensearch-service>
- **Elasticsearch Guide v8.17:** <https://www.elastic.co/guide/en/elasticsearch/reference/8.17/index.html>

© Blend 2026. Todos los derechos reservados.