

langgenius / dify

453 Branches 169 Tags

Go to file Code

thegdsk refactor: migrate _model_to_insertion_dict return type to TypedDict ... 08:25:57 · 12 hours ago 10,068 Commits

.agents/skills	refactor(web): migrate remaining toast usage (#34433)	2 weeks ago
.claude	refactor(web): consolidate query/mutation guidance and dep...	last month
.devcontainer	feat(tasks): isolate summary generation to dedicated dataset...	last month
.gemini	chore: Add initial configuration for Gemini (#34012)	3 weeks ago
.github	docs(contributing): move agent attribution guidance to PR te...	2 days ago
.vite-hooks	fix(ci): repair i18n bridge and translation workflow (#34738)	4 days ago
.vscode	feat(tasks): isolate summary generation to dedicated dataset...	last month
api	refactor: migrate _model_to_insertion_dict return type to Typ...	12 hours ago
dev	fix: update how ky handle error (#34735)	4 days ago
docker	fix(docker): restore S3_ADDRESS_STYLE env examples (#...	4 days ago
docs	Docs: unify language switch links across root and localized ...	2 weeks ago
e2e	test: add e2e scenarios for app creation and sign-out (#34285)	4 days ago
images	feat: update banner (#23095)	9 months ago
packages/iconify-collections	refactor: split icon collections (#34453)	2 weeks ago
scripts/stress-test	refactor: deduplicate legacy section mapping in ConfigHelp...	last month
sdfs	chore: remove raw vite deps (#34726)	5 days ago
web	refactor: remove base ui i18n dependency (#34921)	yesterday
.coveragerc	test: Consolidate API CI test runner (#29440)	4 months ago

About

Production-ready platform for agentic workflow development.

dify.ai

python agent workflow automation ai mcp nextjs orchestration gemini openai gpt low-code no-code rag gpt-4 i18n general agentic-framework agentic-ai

Readme

View license

Code of conduct

Contributing

Activity

Custom properties

137k stars

781 watching

21.5k forks

Report repository

Releases 161

v1.13.3 (Latest) 2 weeks ago

+ 160 releases

Contributors 1,247

Dify.ai repo

Dify es una plataforma de desarrollo de aplicaciones de IA generativa diseñada para ingenieros de software, equipos de producto y departamentos de innovación. Permite orquestar flujos de trabajo complejos, sistemas RAG y agentes autónomos mediante una interfaz visual intuitiva. Es la herramienta ideal para implementar soluciones basadas en modelos de lenguaje (LLM) sin gestionar infraestructuras complejas, permitiendo el despliegue tanto en la nube como en servidores propios mediante Docker.

[Visitar Sitio Oficial](#) [Preguntar a ChatGPT](#) [Preguntar a Claude](#) [Preguntar a Grok](#)

Contenido del Dossier

- [Información de la Herramienta](#)
- [Consejos de Implantación](#)
- [Preguntas Frecuentes](#)
- [Contratos y Condiciones](#)

INFORMACIÓN DE LA HERRAMIENTA

Qué y para quién es

Dify es una plataforma de desarrollo de aplicaciones de IA generativa (LLMOps) que combina la facilidad de las herramientas "no-code" con la potencia de flujos de trabajo avanzados. Permite a profesionales y empresas crear, operar y mejorar aplicaciones basadas en modelos de lenguaje (LLM) de forma visual. Está diseñada para ingenieros de software, analistas de datos, equipos de producto y departamentos de innovación que buscan implementar soluciones de IA (como agentes autónomos o sistemas RAG) sin gestionar infraestructuras complejas desde cero.

Principal ventaja profesional

La capacidad de orquestar flujos de trabajo complejos (Workflows) y sistemas de recuperación de información (RAG) mediante una interfaz visual intuitiva, manteniendo el control total sobre los datos y permitiendo el despliegue tanto en la nube como en servidores propios (On-premise).

Para quién no es

No es para usuarios finales que buscan un simple chat de IA (como ChatGPT) sin intención de construir herramientas propias. Tampoco es ideal para empresas que prohíben estrictamente el uso de contenedores (Docker) o que no tienen capacidad mínima para gestionar una infraestructura técnica básica en caso de optar por la versión auto-alojada.

Funcionalidades clave

- Orquestación visual de flujos de trabajo (Workflow): creación de lógica compleja mediante nodos.
- Pipeline RAG nativo: gestión de documentos, limpieza de datos y almacenamiento en bases de datos vectoriales integrado.
- Agentes autónomos: soporte para agentes que utilizan herramientas y toman decisiones basadas en objetivos.
- Catálogo de modelos: integración nativa con más de 20 proveedores (OpenAI, Anthropic, Llama 3, Hugging Face, etc.).
- Backend como servicio (BaaS): generación automática de APIs para integrar la IA en aplicaciones externas de forma inmediata.
- Observabilidad y logs: seguimiento detallado de ejecuciones, costes y rendimiento de los modelos.

Precios

- Versión gratuita (Open Source / Sandbox):
- Auto-alojado (Docker): Gratis y completa (sin límites de mensajes, gestionas tu propia infraestructura).
- Cloud Sandbox: Gratis limitada (200 créditos de mensaje, 1 equipo, 5 aplicaciones).
- Rango de precios Cloud (59\$ - 159\$ mes):
- Professional: 59\$/mes (5.000 créditos, 3 miembros, 50 aplicaciones).
- Team: 159\$/mes (10.000 créditos, 50 miembros, aplicaciones ilimitadas).
- Enterprise: Consultar precio para soluciones personalizadas y soporte dedicado.

Perfil del usuario

- Empresas tecnológicas y startups que desarrollan productos con IA integrada.
- Departamentos de IT que necesitan centralizar el acceso a modelos de IA de forma segura.
- Consultoras de transformación digital que implementan soluciones de automatización para clientes.
- Perfiles profesionales: Desarrolladores Full-stack, Ingenieros de IA/ML, Product Managers técnicos y Arquitectos de Soluciones.

Nivel técnico requerido

- Uso: Bajo-Medio (La interfaz visual facilita el uso a perfiles de producto).
- Instalación/Configuración: Medio-Alto (Requiere conocimientos de Docker y gestión de servidores Linux para la versión auto-alojada).
- Necesidades de soporte: Departamentos de sistemas (DevOps) para el mantenimiento de la infraestructura y bases de datos.
- Conocimientos necesarios: Comprensión básica de prompts, gestión de APIs y conceptos de bases de datos vectoriales.

Ejemplos de uso profesional

- Atención al cliente inteligente: Bots que consultan manuales técnicos internos para responder dudas complejas.
- Automatización de procesos: Flujos que extraen datos de correos electrónicos y los introducen en un CRM.
- Asistentes de redacción: Herramientas para departamentos de marketing adaptadas al tono de voz de la marca.
- Análisis de documentos: Sistemas que resumen y comparan contratos legales o informes financieros masivos.

Uso y distribución

- Versión web (Cloud gestionado por Dify).
- Versión escritorio (Docker local en Windows, Mac o Linux).
- CLI (Línea de comandos para despliegues técnicos).

Open source

Dify es de código abierto bajo licencia Apache 2.0, permitiendo su modificación y despliegue privado sin costes de licencia de software.

Integraciones

- Facilidad de integración: Low-code / Full-code.
- API propia: Dispone de una API robusta para consumir las aplicaciones creadas desde cualquier sistema externo.
- Integraciones nativas: Conexión directa con Notion, Google Search, herramientas de base de datos y múltiples proveedores de bases de datos vectoriales (Weaviate, Pinecone, etc.).

Notas finales

Información legal, licencias y contratos

El software se distribuye principalmente bajo licencia Apache 2.0 en su repositorio de GitHub. Para la versión Cloud, el contrato se rige por los términos de servicio de LangGenius, Inc., con cumplimiento de estándares de seguridad como SOC Type II en planes superiores.

Para más información:

- Sitio web oficial: <https://dify.ai>
- Precios: <https://dify.ai/pricing>
- Documentación técnica: <https://docs.dify.ai>
- Github: <https://github.com/langgenius/dify>
- Discord: <https://discord.gg/fD8p3SAt9F>

CONSEJOS DE IMPLANTACIÓN

Aplicación profesional

- Empresas tecnológicas, consultoras de transformación digital y departamentos de I+D que requieren centralizar el desarrollo de soluciones basadas en LLM (Large Language Models).
- Presupuesto: Desde gratuito (opción auto-alojada en servidor propio) hasta planes corporativos de 159\$/mes o superiores para servicios en la nube.
- Puntos clave: Reducción drástica del tiempo de comercialización (time-to-market), estandarización de flujos de trabajo de IA y soberanía de datos mediante despliegue on-premise.

Madurez digital requerida

- Usuarios: Es necesario un equipo con capacidad de diseño lógico y comprensión de ingeniería de prompts. No requiere programación avanzada para el uso diario, pero sí autonomía técnica.
- Empresa: Organizaciones con una infraestructura de datos definida o en proceso de digitalización, que posean una estrategia clara sobre el uso de la Inteligencia Artificial y la seguridad de la información.

Plan orientativo de implantación

Pasos necesarios y estimaciones

- Tiempos estimados de despliegue: De 1 a 3 semanas para una versión operativa inicial en entorno corporativo.
- Evaluación inicial (Semana 1): Definición de casos de uso (RAG, Agentes o Workflows), inventario de APIs de modelos necesarias (OpenAI, Anthropic, etc.) y elección del método de despliegue (Cloud o Docker).
- Configuración inicial (3-5 días): Instalación de la instancia en servidores Linux o configuración del espacio de trabajo en la nube. Conexión de bases de datos vectoriales y fuentes de conocimiento (PDFs, Notion, bases de datos SQL).
- Prueba de concepto (1 semana): Creación de un primer agente o flujo de trabajo para un departamento específico (ej. Soporte Técnico o Ventas). Ajuste de hiperparámetros y evaluación de la calidad de respuesta.
- Formación y despliegue (3-5 días): Capacitación a los administradores de la plataforma y usuarios clave en la creación de prompts y gestión de logs.

Necesidades de formación del equipo

- Fundamentos de IA Generativa y arquitectura RAG (Retrieval-Augmented Generation).
- Gestión de flujos de trabajo lógicos y manejo de variables dentro de Dify.
- Administración de contenedores Docker (solo para la versión auto-alojada).
- Seguridad y gobernanza de datos aplicada a modelos de lenguaje.

Perfiles necesarios

- Perfiles técnicos: Ingeniero de DevOps para el despliegue y mantenimiento de la infraestructura. Desarrolladores de Backend para la integración de la API de Dify con sistemas existentes.
- Personal externo recomendado: Consultores expertos en IA Generativa para la arquitectura de prompts complejos y optimización de flujos de trabajo.
- Otros: Líderes de producto (Product Owners) para definir la lógica de negocio que seguirá la IA.

Retorno de la inversión (ROI)

- Tiempos: Se estima una recuperación de la inversión en 3 a 6 meses mediante la automatización de tareas repetitivas y la mejora en la eficiencia de búsqueda de información interna.
- Cómo medirlo (KPIs): Reducción del tiempo de respuesta en atención al cliente, disminución del error humano en la extracción de datos documentales, número de procesos automatizados con IA y ahorro en costes de desarrollo de software a medida.

Otros

- Escalabilidad: Dify permite escalar horizontalmente mediante contenedores, lo que facilita el crecimiento según aumente la demanda de peticiones de IA en la empresa.
- Privacidad: Al permitir el auto-alojamiento, es una solución idónea para cumplir con normativas estrictas de protección de datos (GDPR) al evitar que la información sensible salga de la red corporativa durante el procesamiento intermedio.

PREGUNTAS FRECUENTES

¿Qué es Dify y en qué se diferencia de un simple chat de IA?

Dify es una plataforma de operaciones de modelos de lenguaje (LLMOps) diseñada para el desarrollo y gestión de aplicaciones de IA generativa. A diferencia de interfaces de chat convencionales, Dify ofrece un entorno estructurado para orquestar flujos de trabajo (workflows), gestionar sistemas de recuperación de información (RAG) y desplegar agentes autónomos a través de una interfaz visual o APIs, permitiendo a las organizaciones construir herramientas personalizadas sobre diversos modelos de lenguaje.

¿Es Dify una solución de código abierto y dónde se puede obtener?

Sí, Dify es una tecnología de código abierto distribuida bajo la licencia Apache 2.0. El código fuente completo, junto con las instrucciones para su despliegue y contribución, está disponible públicamente en su repositorio oficial de GitHub gestionado por LangGenius.

¿Cuáles son las opciones de despliegue y cómo afectan a la soberanía de los datos?

Dify permite dos modalidades principales: una versión Cloud (SaaS) gestionada por el proveedor y una versión auto-alojada mediante contenedores Docker. La opción auto-alojada es especialmente relevante para profesionales que requieren control total sobre su infraestructura y cumplimiento estricto de privacidad, ya que permite mantener los datos dentro de servidores propios o nubes privadas (On-premise).

¿Cómo aborda Dify la seguridad y el cumplimiento normativo?

La plataforma está diseñada con estándares de seguridad industriales, incluyendo el cumplimiento de SOC 2 Type II en sus planes superiores. Al permitir el despliegue local, facilita a las empresas españolas y europeas la alineación con el RGPD (Reglamento General de Protección de Datos), ya que el procesamiento de la información puede realizarse en entornos controlados por la propia organización.

¿Qué costes tiene el uso de la plataforma en entornos profesionales?

Dify ofrece una versión gratuita 'Self-hosted' sin límites de mensajes para quienes gestionan su propia infraestructura. En su versión Cloud, dispone de un Sandbox gratuito limitado y planes de pago (Professional y Team) que oscilan entre los 59\$ y 159\$ mensuales, variando según la capacidad de procesamiento de mensajes, número de aplicaciones y miembros del equipo.

¿Qué tecnologías y modelos de IA integra Dify de forma nativa?

La plataforma incluye un catálogo con más de 20 proveedores de modelos, incluyendo OpenAI, Anthropic, Google (Gemini) y modelos locales como Llama 3 a través de integraciones con Hugging Face u Ollama. Además, se integra de forma nativa con bases de datos vectoriales como Pinecone o Weaviate y herramientas de productividad como Notion o Google Search.

¿Qué nivel de conocimientos técnicos se requiere para operar la herramienta?

Para la creación de aplicaciones y flujos de trabajo, se requiere un nivel medio-bajo gracias a su interfaz visual (Low-code). Sin embargo, para la instalación y mantenimiento de la versión auto-alojada, es necesario un nivel técnico medio-alto con conocimientos sólidos en Docker, gestión de servidores Linux y administración de bases de datos.

¿Qué capacidades ofrece para la implementación de sistemas RAG?

Dify simplifica el pipeline de generación aumentada por recuperación (RAG) al incluir herramientas integradas para la limpieza de datos, segmentación de documentos (chunking) y almacenamiento en bases de datos vectoriales. Esto permite a las aplicaciones consultar información privada de la empresa para generar respuestas contextualizadas y precisas.

¿Es posible integrar Dify con aplicaciones empresariales existentes?

Sí, Dify funciona como un Backend-as-a-Service (BaaS). Cada aplicación creada en la plataforma genera automáticamente una API robusta y documentación propia, permitiendo a los desarrolladores consumir las capacidades de IA desde CRMs, ERPs o aplicaciones web personalizadas sin necesidad de programar la lógica del LLM desde cero.

CONTRATOS Y CONDICIONES

Principales recomendaciones

- **Priorizar el auto-alojamiento (On-premise)**: Para empresas españolas, la instalación mediante Docker en servidores propios (o nubes privadas en la UE) es la opción más segura para garantizar la soberanía de los datos y evitar transferencias internacionales automáticas.
- **Configuración de "Sandbox"**: Al usar código personalizado dentro de los flujos, asegúrese de activar el entorno aislante (Sandbox) para evitar que procesos maliciosos accedan al sistema anfitrión.
- **Gestión de proveedores LLM**: Dify actúa como orquestador. El cumplimiento legal final dependerá de qué modelos conecte (OpenAI, Anthropic, etc.). Debe firmar un DPA (Acuerdo de Procesamiento de Datos) con cada proveedor que integre.
- **Anonimización en RAG**: Antes de subir documentos a la base de conocimientos (Knowledge Base), elimine datos personales sensibles para cumplir con el principio de minimización de datos.

Ley de Inteligencia Artificial (AI Act)

- **Clasificación de riesgo**: Como sistema de orquestación, Dify puede considerarse de "Alto Riesgo" si se usa para propósitos críticos (recursos humanos, infraestructuras, justicia). En estos casos, la empresa debe realizar una evaluación de impacto antes del despliegue.
- **Transparencia (Art. 13)**: Dify permite citar fuentes en sus respuestas. Es obligatorio configurar estas funciones para informar al usuario de que está interactuando con una IA y de dónde proviene la información.
- **Vigilancia humana (Art. 14)**: Se recomienda implementar nodos de "revisión manual" o logs de auditoría técnica dentro de los flujos de trabajo para cumplir con el requisito de supervisión humana efectiva.

Privacidad y protección de datos

- **Responsabilidades**: En la versión Cloud, LangGenius Inc. es el encargado del tratamiento y la empresa española es el responsable del tratamiento. En la versión auto-alojada, la empresa española asume el control total y la responsabilidad directa.
- **Ubicación de los datos (Cloud)**: Los servidores de Dify Cloud se encuentran en la región US-East de AWS (Estados Unidos).
- **Transferencia internacional**: El uso de la versión Cloud implica una transferencia de datos a EE.UU. Aunque Dify cuenta con certificaciones SOC 2 e ISO 27001, las empresas españolas deben verificar que se amparan en las Cláusulas Contractuales Tipo (SCC).
- **Derechos ARCO**: La plataforma permite la eliminación de cuentas y datos asociados, facilitando el ejercicio del derecho de supresión. En la versión auto-alojada, la empresa debe implementar sus propios procedimientos para exportar o borrar datos de usuarios finales.

Propiedad intelectual

- **Propiedad de los datos**: Los términos de servicio especifican que el usuario mantiene todos los derechos sobre los datos de entrada ("Customer Data") y las aplicaciones de IA creadas.
- **Propiedad del resultado**: Los resultados generados por los modelos pertenecen al usuario, pero están sujetos a las licencias de cada proveedor de modelos (ej. términos de OpenAI o Anthropic).

Usos y prohibiciones

- **Usos prohibidos**: Se prohíbe explícitamente el uso para actividades de "Alto Riesgo" según la AI Act sin los controles adecuados, así como el tratamiento de datos personales de categorías especiales (datos biométricos, salud, etc.) en la versión estándar.
- **Usos admitidos**: Automatización de flujos de trabajo profesionales, sistemas de consulta documental interna (RAG) y creación de agentes de soporte siempre que se declare su naturaleza artificial.

Seguridad y certificaciones

- **Seguridad**: Cifrado de claves API en reposo. En la versión Cloud, los secretos están protegidos para que ni siquiera el equipo de Dify pueda verlos.
- **Certificaciones**: Dify (LangGenius Inc.) cuenta con certificaciones SOC 2 Tipo I y Tipo II, además de ISO 27001:2022.

Otros

- **Licencia de Software**: La versión de código abierto utiliza la licencia **Apache 2.0**, lo que permite su uso comercial sin costes de licencia, siempre que se respeten los avisos de autoría originales.
- **Identidad corporativa**: LangGenius, Inc. es una empresa registrada en Delaware, EE.UU., con sede en

Mountain View, California.

Fuentes consultada:

- [Términos de Servicio \(Dify.ai\)](#)
- [Política de Privacidad \(Dify.ai\)](#)
- [Guía de cumplimiento EU AI Act \(Github Oficial\)](#)
- [Política de Cookies](#)
- [Documentación de seguridad y reportes SOC2/ISO](#)
- [Documentación sobre despliegue Cloud vs Self-hosted](#)

Para más información y herramientas:

Explora look4.tools para descubrir las mejores soluciones tecnológicas del mercado.

[Inicio](#) [Todas las herramientas](#) [Categorías](#)

Este documento ofrece recomendaciones generadas mediante análisis humano y sistemas de IA automatizados. La información tiene carácter meramente informativo y no constituye asesoramiento legal, profesional ni garantía de resultados. Las marcas, logotipos y nombres comerciales pertenecen a sus respectivos propietarios y se utilizan únicamente con fines identificativos.