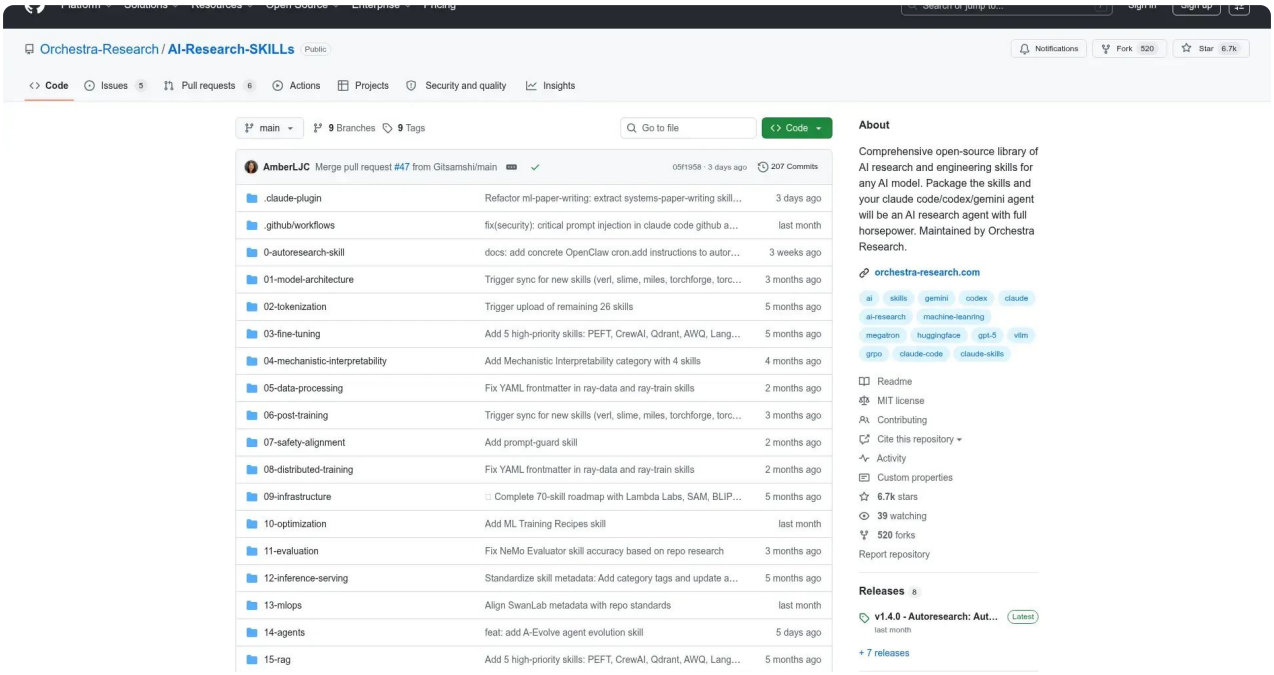

AI Research SKILLs

Biblioteca de código abierto diseñada para ingenieros de ML e investigadores de IA que buscan automatizar el ciclo de vida de la investigación científica. Permite que agentes como Claude Code o Cursor ejecuten tareas de ingeniería complejas, desde la prospección de literatura y generación de ideas hasta el entrenamiento de modelos (RLHF, cuantización) y la redacción automática de artículos en LaTeX. Es la herramienta definitiva para laboratorios de I+D que requieren acelerar su experimentación técnica.

Visitar Sitio Oficial Preguntar a ChatGPT Preguntar a Claude Preguntar a Grok

Contenido del Dossier

INFORMACIÓN DE LA HERRAMIENTA

Qué y para quién es
AI Research SKILLs es una biblioteca de código abierto que proporciona un conjunto exhaustivo de habilidades técnicas y de ingeniería diseñadas específicamente para agentes de IA (como Claude Code, Cursor o Gemini). Su propósito es permitir que estos agentes realicen investigaciones de aprendizaje automático (Machine Learning) de forma autónoma, abarcando todo el ciclo de vida: desde la prospección de literatura y generación de ideas hasta la ejecución de experimentos técnicos y la redacción de artículos científicos en LaTeX.

Está dirigido a ingenieros de ML, investigadores de IA y desarrolladores de software que buscan automatizar las tareas de infraestructura y experimentación para centrarse en la validez científica de sus hipótesis. Es ideal para departamentos de I+D y laboratorios tecnológicos que operan con ciclos rápidos de prototipado de modelos.

Principal ventaja profesional
Permite delegar la configuración técnica compleja (Fine-tuning, optimización distribuida, cuantización) a agentes de IA mediante "habilidades" pre-empaquetadas y verificadas, reduciendo drásticamente el tiempo dedicado a depurar infraestructura y acelerando la producción de resultados publicables.

Para quién no es
No es adecuado para profesionales que buscan una herramienta de análisis de datos generalista o para usuarios sin conocimientos sólidos en Machine Learning. Tampoco es apto para empresas que no utilicen agentes de codificación avanzados (AI coding agents) en su flujo de trabajo o que operen en entornos donde la autonomía de la IA en la ejecución de código esté estrictamente restringida por políticas de seguridad.

Funcionalidades clave
- Autoresearch: Orquestación autónoma de la investigación mediante una arquitectura de doble bucle (optimización interna + síntesis externa).
- 87 Habilidades integradas: Cubre 22 categorías técnicas que incluyen arquitectura de modelos, post-entrenamiento (RLHF, GRPO), inferencia y serving.
- ML Paper Writing: Automatización de la redacción académica con plantillas LaTeX y verificación de citas.
- Gestión de Infraestructura: Soporte para despliegues en la nube (Modal, SkyPilot, Lambda Labs).
- Herramientas de Optimización: Implementaciones de Flash Attention, cuantización (GPTQ, AWQ, GGUF) y entrenamiento distribuido (DeepSpeed, FSDP2).

Precios
- Versión gratuita: La herramienta es Open Source bajo licencia MIT, disponible de forma completa y gratuita en su repositorio de GitHub.
- Rango de precios: 0€ (Licencia gratuita). Los costes asociados derivarán del consumo de computación (GPUs) y el uso de las APIs de los modelos de lenguaje (Claude, OpenAI, etc.).

Perfil del usuario
- Ingenieros de Machine Learning (MLE).
- Investigadores de IA (AI Researchers).
- Desarrolladores de MLOps.
- Departamentos de Innovación y Centros de investigación académica.

Nivel técnico requerido
- Nivel técnico de uso: Alto. Se requiere comprensión profunda de conceptos de IA y manejo de agentes de codificación CLI.
- Nivel técnico de instalación: Medio. Instalación mediante gestores de paquetes (npm/npx) y configuración de entornos de agentes.
- Necesidades de soporte: Equipos de ingeniería para la gestión de infraestructura cloud y cuotas de cómputo.
- Competencias necesarias: Python, manejo de entornos virtuales, conocimientos en frameworks de ML (PyTorch, Hugging Face) y manejo de terminal/CLI.

Ejemplos de uso profesional
- Automatización del proceso de fine-tuning de un modelo de lenguaje específico utilizando técnicas PEFT

o Unsloth.
- Generación autónoma de una comparativa de rendimiento entre diferentes arquitecturas de modelos (Mamba vs Transformers).
- Redacción automatizada de la sección de metodología y resultados de un "paper" técnico tras la ejecución de experimentos.
- Optimización de la inferencia de modelos para su despliegue en entornos de producción mediante cuantización y vLLM.

Uso y distribución
- Versión web: Documentación oficial y portal de bienvenida.
- CLI: Herramienta de instalación interactiva vía npx.
- Extensiones/Plugins: Marketplace de Claude Code y compatibilidad con agentes como Cursor, OpenClaw, Codex y Gemini CLI.

Open source
Distribuido bajo licencia MIT, permitiendo el uso comercial, modificación y distribución privada o pública.

Integraciones
- Facilidad de integración: Nivel técnico medio-alto (requiere configuración de plugins en agentes).
- API propia: No dispone de una API de servicio, funciona como una biblioteca de prompts y scripts estructurados para agentes.
- Integraciones nativas: Compatible con frameworks líderes como DeepSpeed, vLLM, Megatron-Core, LangChain y LlamaIndex.
- Ejemplos de integración: Conexión con Weights & Biases (W&B) para observabilidad y Modal para ejecución serverless de GPUs.

Notas finales
Información legal, licencias, contratos
El proyecto está licenciado bajo MIT License. El software se proporciona "tal cual", sin garantías de ningún tipo por parte de Orchestra Research. La propiedad intelectual de los resultados generados (papers, código) pertenece al usuario, según los términos estándar de la licencia MIT.

Para más información:
- Sitio web oficial: https://www.orchestra-research.com/ai-research-skills
- Github: https://github.com/Orchestra-Research/AI-research-SKILLs
- Linkedin: https://www.linkedin.com/company/orchestra-research/
- Twitter / X: https://x.com/orch_research

CONSEJOS DE IMPLANTACIÓN

Aplicación profesional
AI Research SKILLs es una biblioteca de ingeniería de alta especialización diseñada para **laboratorios de I+D, departamentos de innovación y centros académicos** de computación. Su uso se centra en automatizar el "trabajo sucio" de la investigación en IA (depuración de infraestructura, configuración de entornos distribuidos y redacción técnica en LaTeX).
- **Tipos de empresa:** Startups de IA generativa, laboratorios de ML (Machine Learning) y equipos de MLOps que trabajan con modelos propios (Large Language Models).
- **Puntos clave:** Optimización de ciclos de experimentación mediante una arquitectura de doble bucle (optimización interna de experimentos y síntesis externa de resultados).
- **Presupuesto:** El software es gratuito (Open Source), pero requiere un presupuesto operativo para **computación en la nube (GPUs de gama alta)** y **créditos de API** para los agentes de IA (Claude, OpenAI o Gemini).

Madurez digital requerida
- **Usuarios:** El equipo debe contar con conocimientos avanzados en **Python, PyTorch, Hugging Face** y manejo fluido de terminal/CLI. Es indispensable la familiaridad con agentes de codificación (como Claude Code o Cursor).
- **Empresa:** Requiere una infraestructura capaz de soportar ejecución de código autónomo por IA y protocolos de seguridad que permitan el despliegue de agentes en entornos de desarrollo.

Plan orientativo de implantación
Pasos necesarios y estimaciones
El despligue técnico es rápido (minutos), pero la integración en el flujo de trabajo científico requiere una fase de validación de 2 a 4 semanas.
- **Configuración inicial (1 día):** Instalación interactiva mediante el comando npx @orchestra-research/ai-research-skills y vinculación con los agentes de codificación existentes (Claude Code, Gemini CLI, etc.).
- **Prueba de concepto (1 semana):** Ejecución de una tarea de investigación acotada, por ejemplo, un fine-tuning específico con **Unsloth** o **PEFT** para validar la conexión con la infraestructura de GPU (Modal, SkyPilot).
- **Configuración del orquestador (3-5 días):** Activación del componente **Autoresearch** para gestionar ciclos continuos de experimentación y revisión de literatura.
- **Capacitación (Continua):** Adaptación de los prompts de investigación y supervisión humana de la validez científica de los resultados generados por el agente.

Necesidades de formación del equipo
El personal debe ser formado específicamente en la **interpretación de los "pensamientos" del agente** y en la supervisión de las habilidades técnicas instaladas (87 habilidades en 22 categorías). Es vital entender el funcionamiento de los marcos de entrenamiento distribuido como **DeepSpeed** y **FSDP2** para corregir posibles errores de configuración del agente.

Perfiles necesarios
- **Perfiles técnicos:** Ingenieros de Machine Learning (MLE), especialistas en MLOps e investigadores científicos (PhD o MS en IA).
- **Personal externo:** Consultores en infraestructura Cloud (si no se tiene experiencia en AWS/GCP/Lambda Labs) para optimizar costes de GPU.

Retorno de la inversión (ROI)
- **Tiempos:** Reducción estimada del **40% al 60%** en el tiempo de configuración de experimentos y redacción de borradores técnicos.
- **KPIs:** Número de experimentos completados por mes, reducción de errores de infraestructura (logs de error), tiempo transcurrido desde la hipótesis hasta el borrador del paper en LaTeX.

Otros
- **Seguridad:** Dado que la herramienta permite que los agentes ejecuten código de forma autónoma, se recomienda el uso de sandboxes o entornos aislados para evitar ejecuciones accidentales en producción.
- **Compatibilidad:** Es compatible con frameworks de vanguardia como **vLLM** para inferencia, **W&B** para observabilidad y **Megatron-Core** para modelos de gran escala.

PREGUNTAS FRECUENTES

¿Qué es exactamente AI Research SKILLs?
Es una biblioteca técnica de código abierto diseñada para dotar a agentes de inteligencia artificial (como Claude Code o Cursor) de capacidades específicas de ingeniería de aprendizaje automático. Su función es actuar como un puente entre el agente de IA y las tareas complejas de un laboratorio de investigación, permitiendo la automatización de todo el ciclo de investigación, desde la lectura de literatura científica hasta la ejecución de experimentos y el redactado de artículos en LaTeX.

¿Para qué sirve en un entorno profesional?
Sirve para automatizar la infraestructura y los procesos repetitivos en departamentos de I+D. Permite delegar en agentes de IA tareas como el ajuste fino (fine-tuning) de modelos, la implementación de técnicas de optimización como cuantización, y la gestión de experimentos técnicos, lo que permite a los investigadores centrarse en el análisis de resultados y la definición de hipótesis científicas.

¿Cuál es el coste de uso del software?
La herramienta es completamente gratuita y se distribuye bajo la licencia MIT. No obstante, el usuario debe considerar los costes operativos indirectos, que incluyen el consumo de computación en la nube (GPUs) a través de proveedores como Lambda Labs o Modal, y el gasto derivado del uso de las APIs de los modelos de lenguaje (LLMs) que alimentan a los agentes.

¿Es open source?
Sí, el proyecto está disponible de forma abierta y su código fuente puede ser auditado, modificado y descargado directamente desde su repositorio oficial en GitHub. Utiliza una de las licencias más permisivas de la industria tecnológica.

¿Cumple con la normativa española de protección de datos?
Al ser una biblioteca de herramientas que se ejecuta en el entorno controlado por el usuario o en su infraestructura contratada, el cumplimiento del RGPD depende de cómo el profesional configure el agente de IA y los servicios de computación de terceros. El software en sí no actúa como un proveedor de servicios en la nube (SaaS) que almacene datos personales en servidores propios.

¿Cómo afronta la privacidad de los datos de investigación?
La privacidad depende estrictamente del entorno donde se despliegue. Al ser una biblioteca que el usuario instala localmente o en su propia nube (vía CLI), los datos de investigación no se comparten con los creadores de AI Research SKILLs. El riesgo de privacidad reside fundamentalmente en las políticas de los modelos de lenguaje externos utilizados (como los de OpenAI o Anthropic).

¿Es una tecnología segura para una infraestructura corporativa?
Es una herramienta técnica avanzada que requiere una gestión cuidadosa de los permisos. Dado que permite a agentes de IA ejecutar código de forma autónoma para realizar investigaciones, no se recomienda su uso en entornos corporativos donde la ejecución de scripts externos no esté supervisada por perfiles técnicos senior con conocimientos en seguridad informática.

¿Qué nivel técnico se requiere para su implementación?
El nivel requerido es alto. El usuario profesional debe tener experiencia en el manejo de terminal (CLI), lenguajes de programación como Python, y conocimientos profundos en frameworks de aprendizaje automático como PyTorch o Hugging Face. No es una solución 'plug-and-play' para usuarios finales sin base técnica.

¿Con qué herramientas y plataformas se integra?
Ofrece integraciones nativas con marcos de trabajo de alto rendimiento como DeepSpeed y vLLM, plataformas de observabilidad como Weights & Biases (W&B), y proveedores de infraestructura serverless como Modal. Es compatible con los principales agentes de codificación del mercado, incluyendo Gemini CLI y OpenClaw.

¿Qué tipo de soporte técnico ofrece?
Al ser un proyecto de código abierto, el soporte principal se gestiona a través de la comunidad en GitHub (mediante 'Issues' y discusiones). No existe un contrato de nivel de servicio (SLA) para usuarios gratuitos, por lo que las organizaciones suelen requerir un equipo de ingeniería interno para su mantenimiento y despliegue.

CONTRATOS Y CONDICIONES

Informe técnico descriptivo

Principales recomendaciones
- **Validación humana obligatoria**: Dado que la herramienta genera artículos científicos y ejecuta experimentos de forma autónoma, los resultados (especialmente citas bibliográficas y códigos de entrenamiento) deben ser revisados por expertos para evitar "alucinaciones" o errores técnicos.
- **Control de ejecución de código**: AI Research SKILLs permite a los agentes ejecutar código en el entorno local o cloud. Se recomienda usar entornos aislados (Docker o máquinas virtuales) para evitar que un error en el script generado comprometa la infraestructura de la empresa.
- **Gestión de costes de terceros**: La biblioteca es gratuita, pero su uso intensivo con modelos como Claude 3.5 o GPT-4, sumado al alquiler de GPUs (Modal, Lambda Labs), puede generar costes operativos imprevistos e elevados.
- **Supervisión de dependencias**: Al integrar múltiples librerías de terceros (vLLM, DeepSpeed, etc.), es necesario realizar auditorías de seguridad periódicas sobre las versiones de estas sub-dependencias.

Ley de Inteligencia Artificial (AI Act)
- **Clasificación**: Generalmente se considera un sistema de IA de **propósito general** o una herramienta de soporte a la investigación. No entra, en principio, en la categoría de "alto riesgo" (Anexo III) a menos que se use para fines específicos como la evaluación de personas o infraestructuras críticas.
- **Transparencia**: El usuario debe informar claramente si los resultados de una investigación o un "paper" han sido generados o asistidos por esta IA, cumpliendo con las obligaciones de transparencia del AI Act.
- **Uso prohibido**: No debe utilizarse para la creación de contenido que explote vulnerabilidades de grupos específicos o para sistemas de puntuación social.

Privacidad y protección de datos
- **Responsabilidades**: La empresa usuaria actúa como **Responsable del Tratamiento**. Orchestra Research, al proporcionar código abierto (MIT), no accede a los datos a menos que se use su plataforma web específica.
- **Ubicación de los datos**: Al ser una herramienta ejecutada principalmente en local o en nubes elegidas por el usuario (Modal, SkyPilot), la ubicación de los datos depende de la infraestructura contratada por la empresa española.
- **Transferencia internacional**: Si el agente de IA utilizado (Claude, OpenAI, Gemini) procesa la información en servidores fuera del Espacio Económico Europeo, se debe formalizar un anexo de transferencia de datos y verificar el cumplimiento del marco de privacidad (Data Privacy Framework).
- **Derechos ARCO**: Al ser una herramienta de investigación técnica, la empresa debe asegurar que los datasets utilizados para el entrenamiento o fine-tuning no contengan datos de carácter personal sin base legal, permitiendo el ejercicio de derechos de supresión u oposición.

Propiedad intelectual
- **Propiedad de datos**: El usuario conserva la propiedad total sobre los datos de entrenamiento y los prompts suministrados localmente.
- **Propiedad del resultado**: Según la licencia MIT y los términos de Orchestra Research, la propiedad intelectual de los resultados generados (código, artículos en LaTeX, modelos optimizados) pertenece al usuario/empresa que opera la herramienta.
- **Licencia de la herramienta**: Distribuido bajo **Licencia MIT**, lo que permite uso comercial, modificación y distribución sin coste de licencia de software, siempre que se mantenga el aviso de copyright original.

Usos y prohibiciones
- **Usos admitidos**: Investigación académica, optimización de modelos de lenguaje, automatización de infraestructuras de ML y redacción técnica asistida.
- **Usos prohibidos**: Desarrollo de armas biológicas, químicas o nucleares, investigaciones destinadas a causar daño personal, generación de contenido fraudulento o intento de ingeniería inversa de los modelos propietarios subyacentes.

Seguridad y certificaciones
- **Seguridad**: La herramienta incluye habilidades específicas de "Safety & Alignment" (LlamaGuard, NeMo Guardrails) para filtrar entradas y salidas peligrosas o no deseadas.
- **Certificaciones**: Al ser un proyecto de código abierto, no cuenta de forma nativa con certificaciones ISO o

SOC2, las cuales deben ser aportadas por el proveedor de infraestructura cloud donde se ejecute (ej. AWS, Google Cloud).

Otros
- **Impacto legal**: **Medio**. Aunque la licencia es permisiva, la autonomía del agente para ejecutar código y realizar gastos en APIs/Cloud requiere una política interna de gobernanza clara.

Fuentes consultadas:
- Repositorio oficial GitHub y términos de licencia
- Términos de servicio de Orchestra Research
- Documentación de contribución y estándares de calidad
- Sitio web oficial del proyecto