

엔터프라이즈 에이전틱 AI 플랫폼: NAMU Agentic AI for Red Hat AI Enterprise

NAMU Agentic AI for Red Hat AI Enterprise - 비용, 복잡성, 역량의 삼중 장벽을 허무는 엔터프라이즈 에이전틱 AI 플랫폼

생성형 AI가 단순 응답형을 넘어 목표 지향적 행동과 멀티스텝 추론, 외부 툴 사용, 자율적 의사결정이 가능한 에이전틱 (Agentic) AI로 진화하면서, 기업은 새로운 패러다임 앞에서 세 가지 벽에 부딪히고 있습니다.

고성능 GPU의 대규모 초기 투자 부담, AI 모델·프레임워크·툴 난립으로 인한 의사결정 피로, 그리고 프로덕션 수준 에이전트 시스템을 운영할 내부 역량의 부재가 도입을 가로막는 핵심 요인입니다.

나무기술과 레드햇, 검증된 인프라 위의 에이전틱 AI

나무기술의 AI 에이전틱 플랫폼 NAA(NAMU Agentic AI)와 Red Hat AI Enterprise (Red Hat OpenShift + Red Hat OpenShift AI)를 긴밀하게 통합한 엔터프라이즈급 에이전틱 AI 솔루션입니다.. 하드웨어부터 에이전트 애플리케이션까지 수직 통합된 풀스택을 제공합니다.



- **엔터프라이즈급 에이전틱 AI:** 챗봇, 보고서 자동 생성, MCP 기반 멀티 에이전트 협업을 플랫폼에서 제공하며, vLLM과 llm-d 기반 분산 서버로 대규모 추론 워크로드를 안정적으로 처리합니다.
- **비용 최적화된 SLM 전략:** SLM과 vLLM 기반 추론 최적화를 통해 클로즈드 소스 API 대비 총소유비용(TCO)을 절감하면서도 엔터프라이즈 품질의 서비스 수준을 확보합니다.
- **유연한 하이브리드 배포:** 온프레미스, 프라이빗 클라우드, 하이브리드 환경 어디서든 동일한 방식으로 AI 에이전트를 배포하고 운영하며, PIPA 등 국내 규제 컴플라이언스 대응을 지원합니다.
- **빠른 가치 실현:** 검증된 참조 아키텍처와 NAA Unified Workflow를 통해 PoC부터 프로덕션 배포까지의 시간을 단축하며, 컨설팅·구현·유지보수를 일괄 제공합니다.

사용 기술

- Red Hat AI Enterprise
- NAA (NAMU Agentic AI)
- vLLM / llm-d 분산 서버
- AI Guardrails & MCP
- LLM & Agent Ops
- GPU/NPU 가속 스택

주요 특징

- 엔터프라이즈 급 통합 풀스택 에이전틱 AI
- SLM 기반 비용 효율적 추론
- 온프레미스·하이브리드 유연 배포
- 국내 규제 컴플라이언스 대응
- 검증된 하드웨어 호환성

기존 AI 인프라와 NAMU Agentic AI for Red Hat AI Enterprise 비교

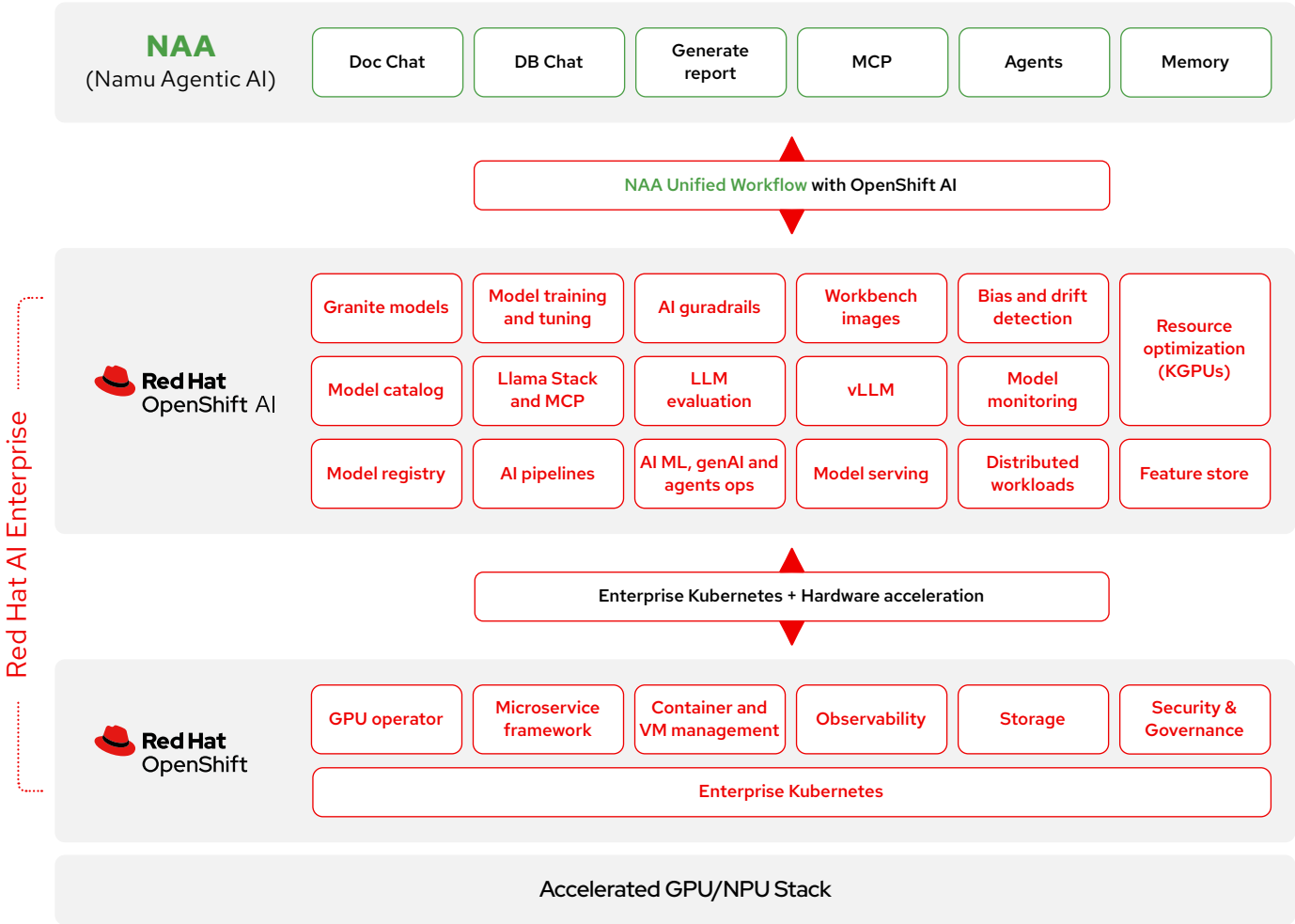
기존 AI 에이전트	NAMU Agentic AI for Red Hat AI Enterprise
고비용 클로즈드 소스 API 의존	SLM + vLLM 기반 비용 효율 추론
AI 모델·프레임워크 선택의 복잡성	검증된 풀스택 LLMOps 제공
분절된 멀티 벤더 공급 체계	단일 벤더 수직 통합 공급
SaaS 전용 또는 온프레미스 제한	온프레미스·하이브리드 동일 방식 배포
에이전트 운영 역량 내재화 부담	에이전트 운영을 위한 LLMOps 포함

솔루션의 주요 기능

- 풀스택 AI 에이전틱 플랫폼:** GPU/NPU 인프라부터 Red Hat OpenShift 애플리케이션 플랫폼, Red Hat OpenShift AI의 모델 서빙·파이프라인·Guardrails 오케스트레이션, 그리고 NAA 애플리케이션 레이어까지 레이어드(layered) 아키텍처로 통합되어 있습니다.
- NAA Unified Workflow:** Doc Chat, DB Chat, Generate Report, MCP, Agents, Memory 등 6대 핵심 기능이 Red Hat OpenShift AI와 긴밀하게 연동되어 기업 업무 자동화와 의사결정 지원을 하나의 워크플로우로 제공합니다.
- Context Engineering과 RAG 파이프라인:** PDF·이미지·텍스트 멀티모달 지식을 문서 추출, 임베딩, 벡터스토어로 처리하고, 단기·장기·시맨틱 메모리와 결합한 SLM 추론으로 정확한 에이전트 응답을 생성합니다.
- 검증된 엔터프라이즈 인프라:** Red Hat OpenShift의 Observability, Security & Governance, Storage, Container/VM 관리 기능을 기반으로 규제 산업에서도 안심하고 운영할 수 있는 안정성을 제공합니다.
- 폭넓은 하드웨어 호환성:** Dell, HP, Lenovo, Nutanix 등 주요 GPU 서버와 NVIDIA, AMD, Rebellions 등 AI 가속 칩셋의 학습 및 추론을 지원합니다.
- 규제 준수 기본 내장:** 개인정보 보호법(PIPA) 등 국내 규제 컴플라이언스가 솔루션 설계에 반영되어 있어 금융, 의료, 공공 등 규제 환경이 엄격한 산업의 요건을 충족합니다.



NAMU Agentic AI for Red Hat AI Enterprise 아키텍처



본 아키텍처는 Red Hat OpenShift AI와 NAA가 유기적으로 결합하여 완성되는 엔터프라이즈 에이전틱 AI 플랫폼을 보여준다. 하단의 하드웨어 가속 레이어는 NVIDIA, AMD, Rebellions 등 다양한 AI 가속기를 지원하며, Red Hat OpenShift가 컨테이너·VM 통합 관리와 거버넌스를 담당한다.

Red Hat OpenShift AI 레이어는 vLLM 기반 모델 서빙, llm-d 분산 추론, AI Guardrails, 데이터사이언스 파이프라인, 모델 레지스트리 등 엔터프라이즈 AI 운영에 필요한 전 기능을 제공하며, 최상단의 NAA는 NAA Unified Workflow를 통해 이들 기능을 단일 에이전트 경험으로 묶어 기업 업무에 즉시 활용 가능한 형태로 제공한다.

산업별 활용 사례 — 바이오·헬스케어

바이오·헬스케어 산업에서는 개인정보 보호와 대규모 연구·임상 데이터 분석이 동시에 요구됩니다. NAA는 이러한 환경에서 규제 준수와 업무 효율성을 함께 충족하는 솔루션으로 활용되고 있습니다.

NAA는 최대 500페이지 분량의 임상 전문 보고서 생성 에이전트, 내부 DB와 연동된 Doc Chat, 보안 에이전트를 제공하여 연구자들이 개인정보 보호 규정을 준수하면서도 논문 및 임상 데이터를 빠르게 분석하고 인사이트를 도출할 수 있도록 지원합니다. 특히 보고서 자동 생성 기능을 통해 문서 작성 시간을 크게 단축하여 연구 생산성 향상에 기여하고 있습니다.

또한 금융 및 제조 산업에서도 NAA는 생성형 보고서, DB Chat, GPU 리소스 최적화 기능을 통해 각 산업의 핵심 Pain Point를 해결하고 있습니다. 금융권에서는 내부 DB 기반의 실시간 지표 분석 및 자동 보고서 생성이 가능하며, 제조 분야에서는 Red Hat OpenShift 기반 GPU 자원 최적화를 통해 운영 비용 절감 효과를 제공하고 있습니다.



간편한 시작

검증된 참조 아키텍처와 사전 구성된 NAMU Agentic AI for Red Hat AI Enterprise 솔루션을 활용하면, 파일럿 에이전트 배포를 수 주 내에 완료할 수 있다.

배포 단계는 다음과 같다.

- **평가 및 계획 (Assessment & planning, 1~2주):** 산업 도메인과 업무 시나리오를 정의하고, 데이터·보안 요건과 배포 아키텍처를 수립한다.
- **스택 배포 (Stack deployment, 1~2 주):** Red Hat OpenShift AI 설치와 NAA Unified Workflow 연동, 지원 하드웨어 통합을 수행한다.
- **파일럿 에이전트 구축 및 검증 (Pilot build & validation, 2~4주):** Doc Chat, Generate Report 등 우선순위 에이전트를 구축하고 실제 업무 데이터로 PoC 검증을 완료한다.

자세히 알아보기

NAMU Agentic AI for Red Hat AI Enterprise는 레드햇 에코시스템 솔루션 Co-Creation 프로그램을 통해 공식화된 엔터프라이즈 에이전틱 AI 플랫폼이다.

바이오·헬스케어, 금융·공공, 제조·Tech 산업에서 PoC 및 도입을 검토 중인 기업은 나무기술 영업부문과 AI부문 담당자를 통해 상세 제안과 데모를 요청할 수 있다.



나무기술 소개

나무기술은 AI·클라우드 분야의 풀스택 전문기업으로, - AI 에이전틱 플랫폼 NAA(NAMU Agentic AI), Edge AI PaaS 솔루션 NEA 등을 자체 개발·공급하고 있다. GPU 서버 인프라부터 AI 추론 엔진 vLLM, 에이전트 오케스트레이터, CMP까지 엔드-투-엔드 AI 시스템을 자체 역량으로 구축·공급할 수 있는 국내 몇 안 되는 기업 중 하나이다.

AI 연구 및 개발, 비즈니스·마케팅 전문 인력을 포함한 팀이 솔루션의 기획부터 구축, 유지보수까지 전 사이클을 책임진다. 국내를 비롯해 싱가포르, 일본, 베트남, 캐나다에서 컨설팅·구현·유지보수 서비스를 제공하며 글로벌 거점을 확장하고 있다. NAMU Agentic AI for Red Hat AI Enterprise의 개발 파트너십을 통해 나무기술은 검증된 인프라 위의 엔터프라이즈 에이전틱 AI라는 명확한 포지셔닝으로 국내외 시장 확대에 나서고 있다.

About Red Hat

Red Hat is the world's leading provider of enterprise open source software solutions, using a community-powered approach to deliver reliable and high-performing Linux, hybrid cloud, container, and Kubernetes technologies. Red Hat helps customers develop cloud-native applications, integrate existing and new IT applications, and automate and manage complex environments. A trusted adviser to the Fortune 500, Red Hat provides award-winning support, training, and consulting services that bring the benefits of open innovation to any industry. Red Hat is a connective hub in a global network of enterprises, partners, and communities, helping organizations grow, transform, and prepare for the digital future.

f facebook.com/redhatinc

X @RedHat

in linkedin.com/company/red-hat

North America
1-888-REDHAT1
www.redhat.com

Europe, Middle East, and Africa
00800 7334 2835
europe@redhat.com

Asia Pacific
+65 6490 4200
apac@redhat.com

Latin America
+54 11 4329 7300
info-latam@redhat.com