

품 목(구매) 규 격 서

품 명	GPU 서버	수 량	1대
		설치장소	한국방사선진흥협회(서울)
		납기일	계약체결일로부터 30일 이내
주요구성 부분 및 SPECIFICATION			
I. 용도(End-user's Use)			
1. 본 연구 장비는 「방사선 기술 융복합 클러스터」 과제의 방사선 연구시설장비 통합정보 플랫폼 (RIIS) 시범 구축·운영을 위한 것으로, 방사선 연구시설·장비 정보와 이에 연계된 논문·특허 등 R&D 성과를 통합 검색·활용하는 검색 특화 AI 모델(임베딩·색인·검색)의 자체 개발·검증에 사용하며, 향후 GPU·저장장치 증설이 가능한 확장형 구성이어야 한다.			
2. 대용량 학습이 아닌 임베딩·색인·경량 검색모델 운용을 목적으로 하는 상시·경량 검색연산 전용 장비로, 단일 GPU 구성으로 공개·소규모 정형 데이터를 협회가 자체적으로 직접 구축·갱신할 수 있어야 하며, 민감·내부 데이터를 외부 유출 없이 자체 처리할 수 있어야 한다.			
3. Linux(Ubuntu 등) 또는 Windows 운영체제 환경에서 TensorFlow, PyTorch 등의 딥러닝 라이브러리와 CUDA 등 GPU 가속 도구를 지원하여, 사용자가 즉시 연구에 활용할 수 있도록 최적화된 개발 환경을 제공해야 한다.			
II. 장비의 구성(Configurations of Goods)			
1. 딥러닝 GPU 워크스테이션(타워형) 1식: CPU, GPU, RAM, 메인보드, SSD, PSU, 냉각 시스템, 케이스 등 포함			
III. 성능 및 규격(Performance and Specification)			
1. 워크스테이션 본체			
1-1) 중앙처리장치(CPU)			
멀티코어 연산을 지원하는 CPU			
코어/스레드: 12코어 이상			
1-2) 그래픽 처리 장치(GPU)			
딥러닝 추론·임베딩 연산에 적합한 GPU 1장(특정 모델 미지정, 최소 요구사양 충족(RTX 5090급 이상))			
GPU 메모리(VRAM): 32GB 이상			
1-3) 메인보드			
상기 CPU·GPU를 안정적으로 지원하는 워크스테이션급 메인보드			
확장성: 향후 GPU·저장장치를 증설할 수 있는 구성(가능 시)			

1-4) 메모리(RAM)

타입: DDR5 (ECC 권장)

용량: 64GB 이상

1-5) 저장장치(SSD)

타입: M.2 NVMe

용량: 1TB 이상(OS·데이터용)

1-6) 파워서플라이(PSU)

정격출력: GPU 안정 구동에 충분한 사무용 정격 전원(220V 이하) 구성

1-7) 냉각 시스템(Cooling System) 및 소음

CPU·GPU 안정 구동을 위한 냉각 구성

CPU·GPU 및 시스템의 과열방지를 위한 냉각 및 온도보호 기능 지원

허용 소음 dB: 40dB 미만

1-8) 케이스(Case)

폼팩터: 타워형(확장 가능한 폼팩터)

1-9) 네트워크

유선 기가비트(1GbE) 이상

2. 소프트웨어

2-1) 운영체제: Linux(Ubuntu 등) 또는 Windows 설치

2-2) GPU 연산환경: CUDA, cuDNN 등 GPU 가속 라이브러리 설치

2-3) 딥러닝 프레임워크: PyTorch, TensorFlow 설치 및 GPU 연동 최적화

2-4) 개발환경: Python, Jupyter 등 개발·실행환경 구성

2-5) 임베딩 모델(오픈소스): 문장 임베딩 모델(sentence-transformers, BGE-E5 등 한국어 지원 모델 포함) 설치·구동

2-6) 벡터데이터베이스(오픈소스): FAISS·Chroma·Qdrant·Milvus·pgvector 중 발주자와 협의하여 1종 이상 설치·구성

2-7) RAG·검색 파이프라인(오픈소스): LangChain 또는 LlamaIndex 기반 임베딩→색인→검색 파이프라인 구성

2-8) 로컬 추론·서빙(오픈소스): Ollama·vLLM·Hugging Face Transformers 등 로컬 추론 환경 구성

2-9) 검증: 설치 완료 후 벡터DB 구축(임베딩·색인) 및 자연어 검색 예제까지 정상 동작 검증

※ 상기 소프트웨어는 모두 오픈소스 기반으로 구성하여 별도 라이선스 비용이 발생하지 않도록 하며, 세부 구성요소·버전은 발주자와 협의하여 확정한다.

- 운영체제는 Linux 또는 Windows로 하며, Windows 제안 시 정품 라이선스 비용을 제안금액에 포함하여야 한다. 단, 그 외 임베딩·벡터DB 등 S/W는 오픈소스 기반으로 구성한다.

IV. 기타 조건(Remark)

1. 본 장비 구축을 위해 H/W, S/W 등 모든 제반 사항을 장비 운용 관리담당 사용자와 협의하여 지정 시간, 지정 장소에서 지원하고 모든 구축 완료 전 시스템 시험, 성능시험, 통합시험 등에 대한 방안을 상세히 제시하여야 한다.
2. 최종 산출물 및 테스트 결과물을 첨부하여 발주자에게 승인 검사 및 테스트를 요청하여야 하며, 승인검사 및 테스트 과정에서 발견된 하자 사항은 만족한 결과를 얻을 때까지 보완, 테스트를 반복적으로 실시해야 한다.
3. 설치 후 문제 발생 시 H/W 및 S/W 통합 기술지원으로 문제를 해결하여야 한다.
4. PSU 안정성 입증을 위한 안전 확인 신고 증명서를 제출한다.
5. 모든 부품은 제조사의 신제품이어야 함.
6. 설치 후 A/S 3년 이상 무상 보증(국내 A/S 지원)
7. 본 구매는 예산 범위 내에서 상기 최소 요구사항을 모두 충족하는 GPU 서버를 공급받는 것을 원칙으로 하며, 공급업체는 제조사·제품명·모델명 및 CPU·GPU·RAM·SSD·PSU 등 세부 구성과 상세 규격이 포함된 제품 제안서를 제출하여야 하고, 해당 제안서의 사양 및 구성은 계약내용에 포함 되는 것으로 본다.
8. H/W 설치에 더하여 상기 소프트웨어(임베딩 모델·벡터데이터베이스·RAG 검색 파이프라인 등 벡터DB화 도구)를 오픈소스 기반으로 설치·구성하고, 벡터DB 구축 및 자연어 검색이 정상 동작함을 검증한 후 인계하여야 하며, 협회 실무자가 자체적으로 운영·갱신할 수 있도록 사용 방법을 함께 안내하여야 한다.
9. 대금 지급조건: 납품 및 검수 후 세금계산서 발행