

과제명	On-Premises 기반 생성형AI 보조교관 체계		
부대명	공군교육사령부 정보통신학교	담당자성함/ 연락처	
예산과제예산	1억 원 이내	예산과제기간	7개월 이내
1. 과제개요	<ul style="list-style-type: none"> <li>• LLM모델을 활용한 맞춤형 교육 플랫폼의 개발 및 적용 가능성을 탐색함으로써, 교육의 질적 향상과 효율적인 학습 환경 구축 방안을 제시하는 데 있음. 기존의 전통적인 교육 방식은 교육사·교육생들을 대상으로 한 획일적인 교육과정을 중심으로 운영되어, 교육생 개개인의 이해도와 학습 속도에 맞춘 세밀한 교육 제공이 어려움. 이에 따라 GPT모델 등을 활용하여 개별 교육생에게 최적화된 맞춤형 학습 환경을 조성하고, 학습자 중심의 교육 방식을 구현.</li> <li>• 주요기능 <ul style="list-style-type: none"> <li>- 교육자료 자동 생성(요약, 슬라이드 초안, 문제 출제)</li> <li>- 질의응답형 AI 교관 기능</li> <li>- 자가학습을 위한 문제 자동 출제 및 상세풀이(피드백) 제공</li> <li>- 사용자 권한 분리 및 사용이력 관리</li> <li>- 폐쇄망 기반 인프라에서의 AI 모델 운용(교육사 내 활용가능한 서버)</li> </ul> </li> </ul>		
2. 과제개발 배경	<ul style="list-style-type: none"> <li>• 교육 현장의 디지털 전환 가속화 <ul style="list-style-type: none"> <li>- 군, 보안기관, 연구기관 등에서도 AI를 활용한 교육 혁신 요구 증가. 특히, 콘텐츠 제작의 반복성과 교관 업무 과중 문제가 대두됨</li> </ul> </li> <li>• 생성형 AI의 교육 활용성 증대 <ul style="list-style-type: none"> <li>- GPT 기반 생성형 AI가 문서 요약, 질의응답, 문제 생성 등 교육적 응용에 탁월. 단, 클라우드 기반 AI는 보안 환경에서 사용이 제한됨</li> </ul> </li> <li>• 온프레미스 환경에서의 AI 도입 필요성 <ul style="list-style-type: none"> <li>- 군/정부기관/고등교육기관 등은 인터넷 미연결 환경에서 독립적 AI 시스템 필요. 외부 유출 우려 없이 내부 데이터를 기반으로 학습·응답 가능한 폐쇄형 AI 인프라 요구됨</li> </ul> </li> <li>• 기존 시스템의 한계 <ul style="list-style-type: none"> <li>- 기존 CBT 기반 교육 시스템은 정적이며, 학습자의 수준에 따른 맞춤형 피드백 제공이 어려움. 정형화된 평가체계에서 탈피해, AI 기반 동적 상호작용 학습환경 필요</li> </ul> </li> </ul>		

<p>3. 기대효과</p>	<ul style="list-style-type: none"> <li>● 국방혁신 4.0 · 공군비전 2050 확산 및 안착과 급변하는 안보환경에 강력하고 스마트한 항공우주력 확보와 스마트 비행단에 적용되고 있는 신기술들에 대한 창의적 문제인식, 해결 제시가 가능한 군사전문형 인재 양성의 기반 마련</li> <li>● 생성형 AI의 도입은 단순한 교육 자동화의 차원을 넘어, 개별 교육생의 필요를 충족시키는 맞춤형 학습, 실전과 유사한 훈련 환경 조성, 학습 데이터 기반의 피드백 및 평가 체계를 강화하는 데 크게 기여할 수 있음. 이는 군사 인력의 전문성과 효율성을 높이고, 나아가 전투력 향상에도 중요한 역할.</li> <li>● 교육 콘텐츠 자동화 및 교관 업무 경감 <ul style="list-style-type: none"> <li>- 교재 요약, 슬라이드 구성, 문제 출제를 자동화함으로써 교관의 교육자료 작성시간을 대폭 단축. 반복적인 질의응답 및 피드백 업무를 AI가 수행하여 교관 본연의 교육 품질 관리에 집중 가능</li> </ul> </li> <li>● 맞춤형 학습 환경 제공 <ul style="list-style-type: none"> <li>- 수강생의 질의에 AI가 실시간 응답하여 개인별 학습 지원.</li> <li>- 서술형 피드백 및 핵심 누락 개념 보완으로 개별 학습효과 강화</li> </ul> </li> <li>● 보안 및 자산 보호 <ul style="list-style-type: none"> <li>- 외부망 없이 온프레미스 환경에서 AI가 작동함으로써 기밀정보 유출 차단</li> <li>- 자체 데이터로 학습되기 때문에 정보 주권 확보 및 교육자산의 내재화 가능</li> </ul> </li> <li>● 기술자립 기반의 AI 인프라 확보 <ul style="list-style-type: none"> <li>- 오픈소스 기반 AI 활용으로 클라우드 서비스 종속성 최소화</li> <li>- 향후 국방, 공공, 연구기관 전반에 확장 가능한 보안형 생성형 AI 모델 확보</li> </ul> </li> <li>● 차세대 스마트 교육 기반 마련 <ul style="list-style-type: none"> <li>- 생성형 AI 기반 상호작용 교육체계 도입으로 미래형 교육혁신 실현</li> <li>- 향후 음성(TTS), 이미지(AI 시각 자료) 등 다중 모달 콘텐츠와의 연계 가능성 확보</li> </ul> </li> </ul>
<p>4. 국내외 기술개발 현황</p>	<ul style="list-style-type: none"> <li>● 국내 <ul style="list-style-type: none"> <li>- 공군은 자체 개발한 국방 GPT를 망분리 환경에서 시범 적용중. (공군GPT, AI번역, AI통합검색, 음성분석, 문서요약)</li> </ul> </li> </ul>



- EBS AI펍톡은 초등학생 영어 회화 연습을 위해 한국어 음성인식 +대화모델을 활용중.
- 세종학당재단은 한국어 특화LLM을 기반으로 외국인 대상 한국어 AI튜터를 활용중.
- 가천대와 부산대는 GPT4모델과 Claude혼합형태인 의학 교육용 AI 가상환자를 활용하여 교육에 적용중.
- 국외
  - 미군의 NIPRGPT는 혁신가와 개발자를 한데 모아 차세대 소프트웨어와 작전 능력을 빠른 속도로 개발하고 공군에 배치하도록 지원하는 체계임. 이 챗봇은 국방부의 공통 액세스 카드(CAC) 인증 시스템과 함께 작동하며 질문에 대답하고 서신, 배경 서류 준비, 프로그래밍 같은 작업을 지원.
  - 미 육군은 CamoGPT라는 생성형 AI도구를 개발하여, 훈련 자료에서 특정 내용을 식별하고 제거하는 작업을 자동화하고 있음. Meta의 LLaMA 3.3모델을 기반으로 개발되었음.

**5. 과제목표**

- 과제 세부목표  
[기능 1] 교육 콘텐츠 자동 생성

항목	상세 설명
기능 개요	교관이 입력한 키워드, 강의안, 문서(PDF 등)를 기반으로 요약, 시각 자료(슬라이드 개요), 문제은행 초안을 자동 생성함
세부 기능	문서 자동 요약 (강의노트, 교안 요약) 슬라이드 개요 초안 생성 개념정리 및 문제 자동 생성 (객관식, 주관식, OX 등) LLM 텍스트 요약: 1)KoAlpaca, GPT-J, MPT-7B 등 기반 Summarization RAG 방식: PDF 학습자료를 인텍싱 후 Embedding 기반 생성 문제 생성: P-Tuning 기반 교육 도메인 특화 질의 템플릿 적용

- [기능 2] 질의응답 기반 보조교관 기능

항목	상세 설명
기능 개요	교관 또는 수강생이 입력한 질문에 대해 강의자료 및 내부 지식 기반으로 즉각 응답하는 AI 챗봇 기능 제공

세부 기능	<p>문서 기반 질의응답 (PDF기반 강의자료 기반)          오픈 질의응답 (사전에 학습된 일반 지식 응답 포함)          누적 질의 이력 기반 개인화 응답 제공          Embedding 기반 QA: 2)FAISS + 3)SentenceTransformer (KoSimCSE 등) 조합으로 문서 검색          LLM 기반 응답: context-aware LLM (4)LLaMA3.3 등)          Streamlit + FastAPI 기반 Web UI 연동</p>
----------	--

[기능 3] 평가(자가학습 성취도평가) 및 피드백 자동화

항목	상세 설명
기능 개요	서술형 및 선택형 평가 응답을 AI가 채점하고 개선 피드백을 자동 제공
세부 기능	<p>수업 후 교육생 자가학습을 위한 성취도 평가 후 핵심 키워드 일치 여부 평가          부족 개념 보완 피드백 자동 생성          유사 예문 및 참고자료 제시          LLM 채점 엔진: 5)LangChain 기반 채점 프롬프트 구성          Keyword Match 기반 자동 평가: 6)TF-IDF 또는 7)BERTScore로 정량 평가          8)KoBART 기반 요약형 피드백 생성</p>

[기능 4] 교육자료 커스터마이징 및 대화형 튜터

항목	상세 설명
기능 개요	교육 주제에 따라 교관이 선택한 스타일로 문체, 표현, 난이도 등을 조정할 콘텐츠 자동 제작성 기능
세부 기능	<p>군사교육, 기술교육 등 분야별 콘텐츠 톤 조절          초급/중급/고급 난이도 조절형 콘텐츠 제공          표/이미지 요약 → 설명형 문장 자동 변환          Prompt Engineering: Few-shot 학습을 통한 사용자 설정 반영          LLM Style Control: ChatGPT-style persona 설정 또는 LoRA 기반 스타일 학습          Pandas + Matplotlib 자동 설명: 데이터 시각화 설명 자동 생성 모듈 연동</p>

[기능 5] 온프레미스 인프라 최적화 및 보안

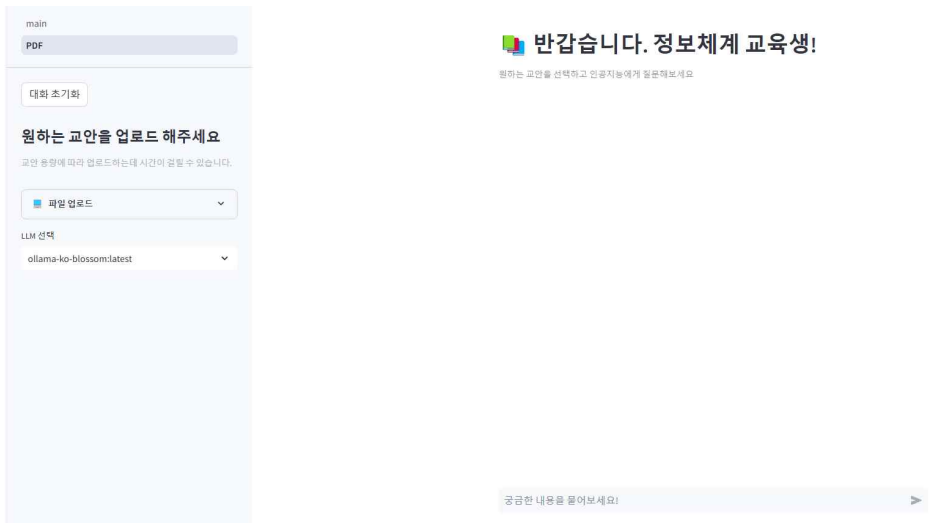
항목	상세 설명
기능 개요	폐쇄망 환경에서 AI 모델과 시스템이 안정적으로 동작하도록 배포/관리 체계 구현
세부 기능	<p>GPU 서버(교육사 서버 활용) 및 모델 경량화(LoRA 등) 지원          클라우드 미사용 환경에서의 AI 서버          Docker + Kubernetes 기반 모델 배포          LoRA, 8-bit quantization 등 모델 경량화 기술</p>

## [기능 6] 사용자 인터페이스 구성

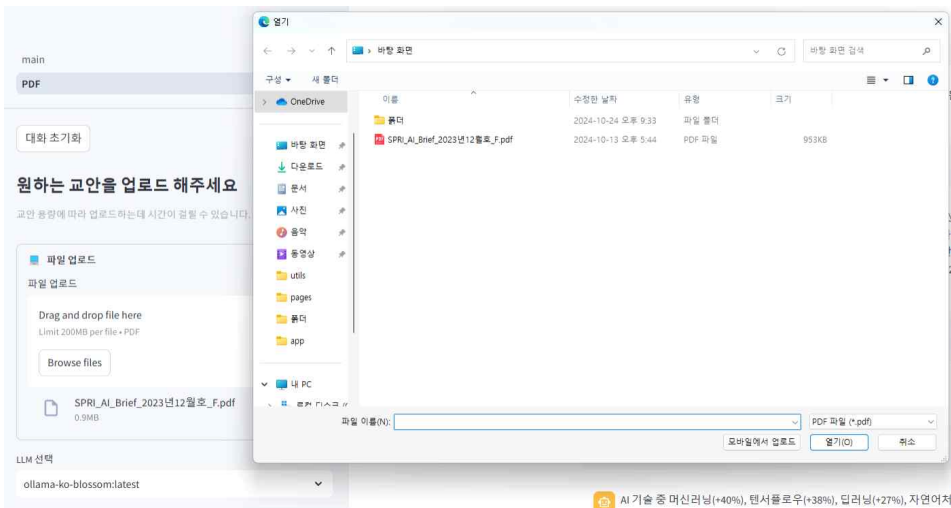
항목	구성 내용
웹 대시보드	교관용 콘텐츠 생성 및 응답 피드백 관리 화면
수강생 대시보드	질문 제출, 피드백 수신, 자료 다운로드
관리자 화면	사용자 현황, 로그 확인, 시스템 설정

### [필수사항]







- 구축 완료된 체계에 군사자료 학습 후 답변을 향상을 위한 Fine-Tuning 필요.(정보통신학교 교육자료)
- 답변수준 : 모든 응답은 GPT-4 수준



### <보조교관 예시화면>



### <PDF기반 교재를 학습시켜 답변을 향상>

	<p> 안녕하세요</p> <p> 안녕하세요! 무엇을 도와드릴까요? 답변이 되었나요</p> <p> 판사 이름을 알려주세요</p> <p> 판사는 윌리엄 오릭(William Orrick)입니다. 답변이 되었나요</p> <p> 목차를 알려주세요</p> <p> 안녕하세요! 저는 SPRI AI Brief의 2023-12월호에 대한 정보를 제공할 수 있습니다. 이 호에는 유튜브가 생성 AI 콘텐츠 표시 의무화, G7이 첨단 AI 시스템 위험 관리 국제 행동강령 마련, 미국 FTC가 소비자 보호와 경쟁 측면의 AI 의견서 제출 등 다양한 주제들이 포함되어 있습니다.</p> <p>목차는 다음과 같습니다:</p> <ol style="list-style-type: none"> <li>1. 정책/법제: 유튜브, 2024년부터 AI 생성 콘텐츠 표시 의무화</li> <li>2. 기업/산업: G7, 히로시마 AI 프로세스를 통해 AI 국제 행동강령 마련</li> <li>3. 기술/연구: 미국 FTC, 저작권청에 소비자 보호와 경쟁 측면의 AI 의견서 제출</li> </ol> <p>이 정보가 도움이 되길 바랍니다! 답변이 되었나요</p> <p style="text-align: center;">&lt;학습된 PDF교재를 바탕으로 답변된 예시&gt;</p>
--	---

- 
- 1) 한국어 명령어를 이해하는 오픈소스 언어모델
  - 2) 밀집 벡터의 유사도 검색과 클러스터링을 위한 라이브러리
  - 3) 문장, 문단, 이미지와 같은 텍스트의 밀집 벡터 표현을 계산하는데 사용되는 프레임워크
  - 4) Meta가 개발한 대규모 언어모델(LLM)
  - 5) LLM을 활용한 어플리케이션을 쉽고 구조적으로 개발할 수 있도록 도와주는 파이썬기반 프레임워크
  - 6) 단어 빈도-역문서 빈도
  - 7) 자연어처리 평가지표
  - 8) SKT에서 공개한 한국어 BART모델