이영주 Young-Ju Lee

데이터·AI 엔지니어

iyju990312@naver.com · github.com/0Ju12 · geoverselab.com/ko/research-history/13/ · LinkedIn

SUMMARY

자원에너지공학을 전공하고 환경에너지융합학과에서 석사 학위를 받았습니다(2026). 학부연구생으로 판구조 형성사를 고려한 해외 대규모 CO2 저장소 탄성파 해석 논문을 제1저자로 썼고, IAB 융합전공 캡스톤에서는 시공간 딥러닝(U-LSTM) 기반 산불 확산 경로 예측 모델의 설계와 학습을 맡았습니다. 석사 과정에서는 중력·자력 데이터의 파수 상관 분석으로 지하 구조를 규명해 SCIE 국제학술지 SPE Journal 제1저자 논문을 냈고, 학부·석사를 합쳐 논문 3편과 학회 발표 6건, 학회 우수논문 발표상의 성과를 냈습니다. 포스코 청년 AI·BigData 아카데미(POSTECH)에서는 FLAN-T5와 RAG를 결합한 영어 표현 교정 시스템 프로젝트에서 데이터 구축과 모델 실험을 맡았고, 팀은 장려상을 받았습니다. 캐나다 캘거리대학교에서 두 차례 파견 연구를 하며 해외 연구진과 협업했고, ADsP(데이터분석 준전문가) 자격을 보유하고 있습니다.

EXPERIENCE · 핵심 활동

석사과정 — 중력·자력 데이터 파수 상관 분석으로 지하 구조 규명 (SPE Journal 제1저자)

2024.03 – 2026.08

연구 · 전북대학교 대학원 환경에너지융합학과 · 지구물리컴퓨팅연구실 · 석사과정 연구원 · 캘거리대 장기 파견

중력·자력 데이터 파수 상관 분석으로 지하 구조 규명 (석사 · SPE Journal 제1저자) 2024.03 – 2026.08 · 석사과정 연구원

상세 ↗

- 상향 연속·정규화로 중력·자력 데이터를 전처리하고, 파수(wavenumber) 영역에서 두 신호의 상관계수 CC(k)를 계산하는 파수 상관 필터링(WCF) 파이프라인 구축
- 양의 상관 성분을 강조해 기반암을 드러내는 SLFI, 약한·음의 상관 성분으로 암염을 드러내는 DLFI 두 지표를 설계
- 기존 3개 독립 모델과 공간 비교·측선별 비교·3D 밀도 모델로 결과를 다각도 검증
- 정기 연구미팅에서 아부다비 자료 처리·분석 결과를 10회 이상 발표(2024.05–2026.01)하고, 2025.11 기초연구실 워크숍에서 연구 성과 발표
- IITP 글로벌 인재양성사업 등 국가R&D 과제 참여 연구원
- 캐나다 캘거리대학교 장기 파견(2024.06–12, 약 6개월): 측지공학과 연구진과 공동 수행, 칼리파대학교(UAE)와 3개 기관 공동연구로 확장
- **SLFI 기반 기반암 구조가 서로 다른 해석법의 기존 독립 모델 3개와 상관계수 0.94 이상으로 일치**
- **DLFI로 암염 구조의 공간 분포를 재현하고, 가샤 서쪽의 추가 암염 융기 가능성을 새로 제시**
- **SCIE 국제학술지 SPE Journal 제1저자 논문 (2026, 캘거리대·칼리파대 공동연구) · 학회 발표 3건 (2024)**
- **2024 KSEG 추계학술대회 학생 우수논문(포스터) 발표상**
- **석사 학위논문 「Characterization of Salt and Basement Structures in the Ghasha Area Using Wavenumber Correlation Analysis of Gravity and Magnetic Anomaly Data」 (전북대학교 대학원 환경에너지융합학과, 2026.08.21 학위 취득)**

캐나다 캘거리대학교 장기 파견연구 — 중력·자력 데이터 복합 해석 2024.06 – 2024.12 · 석사과정 파견연구원 상세 ↗

- 과학기술정보통신부·IITP 디지털분야 글로벌 인재양성사업(AI·디지털트윈 기반 스마트 CCS 운영 글로벌 인재양성 프로그램) 지원으로 파견
- 캘거리대학교 측지공학과 연구진과 함께 항공 중력·자력 데이터의 파수 상관 분석 연구 수행
- 칼리파대학교(UAE)와의 3개 기관 공동연구로 확장
- 캐나다 로키 카나나스키스(Kananaskis) 지역 현장 지구물리 탐사에 참여하고 영문 현장 탐사 보고서(Field Exploration Report) 공동 작성 (2024.08)
- **파견 연구 결과로 SCIE 국제학술지 SPE Journal 제1저자 논문 게재 승인 (2026)**
- **같은 연구를 석사 학위논문(2026)으로 완성**

신호처리(파수 영역) · 상관 분석 · 지표 설계 · 중력·자력 데이터 · 국제공동연구

학부 — 판구조 기반 해외 CO2 저장소 탄성파 해석 (제1저자) · IAB 융합전공 시공간 딥러닝 2022 – 2024.02

연구 · AI 프로젝트 · 전북대학교 자원에너지공학과 · IAB 융합전공 · 지구물리컴퓨팅연구실 · 학부연구생 · 캡스톤 모델 설계·학습 담당

판구조 형성사를 고려한 해외 대규모 CO2 저장소 탄성파 해석 (제1저자 논문) 2022 – 2024.02 · 학부연구생 · 제1저자 상세 ↗

- GPlates(PALEOMAP PaleoAtlas)로 북미·북해·오세아니아·중동 4개 지역의 판구조 변천사를 복원해 CO2 저장소가 만들어진 과정을 해석
- 지역별 대표 CCS 프로젝트(북미 Quest·Weyburn-Midale, 북해 Northern Lights·PORTHOS, 오세아니아 Gorgon·Moomba, 중동 Uthmaniyah·카타르 LNG — 지역별 2개씩)의 탄성파 영상화 결과를 저장층·덮개암 지질구조 측면에서 해석
- 선행 연구: GPU 기반 4D 탄성파 모니터링 기술 적용성 검토 (2022) / 사우디아라비아 딥러닝 기반 탄성파 영상화·노르웨이 슬라이프너 지구물리 모니터링 사례 분석 (2023)
- 캘거리대학교 단기 파견(2023.12–2024.02): 미국 Utah FORGE 지열 부지 중력 데이터 처리
- **제1저자 논문 「판구조론적 역사를 고려한 해외 대규모 이산화탄소 지중저장소 탄성파 해석 결과 분석」 (Analysis on Seismic Interpretation for Overseas Large-scale CO2 Storage Considering Geological History Related to Plate Tectonics), 지구물리와 물리탐사 27(1), 1–22, 2024 — 학부 과정 연구 성과**
- **제2저자 논문 1편 (호주 깁스랜드 분지 CarbonNet CCS 지질학적 분석, 지구물리와 물리탐사, 2023) · 학회 발표 3건 (2022–2023)**
- **각 프로젝트의 탄성파 영상에 드러난 층서구조가 해당 지역의 판구조적 형성 역사를 반영함을 확인 → 국내 탐사 자료 처리 시 지질학적으로 비현실적인 속도 모델을 배제하는 기준으로 활용할 수 있음을 제시**

IAB 융합전공 — 시공간 딥러닝 기반 산불 확산 경로 예측 모델 (캡스톤) 2023.09 – 2023.12 · 모델 설계·학습 담당 상세 ↗

- 입력 래스터: 산림청 산불 상황도를 GIS로 디지타이징한 화선 + 기상·NDVI·고도
- Conv2D 인코더 → LSTM → Conv2DTranspose 디코더 구조의 U-LSTM 모델을 설계하고, 데이터가 적은 점을 고려해 파라미터 2,736개의 얕은 네트워크로 구성
- 부족한 시점 데이터를 길이 1~19의 슬라이딩 윈도우로 확장해 학습 배치 151개 구성
- TensorFlow·PyTorch로 모델을 구현하고 Overall Accuracy·F1-score로 검증·평가
- 입력 데이터: 4개 대형 산불(울진·삼척, 군위, 밀양, 강릉)의 화선·고도·NDVI·풍향·풍속을 ArcGIS로 2D 래스터화 — 3개로 학습, 1개는 미학습 테스트로 일반화 성능 검증
- **미학습 산불(강릉) 테스트 Overall Accuracy 0.90 · F1-score 0.85 (검증 0.95 · 0.81) — 초기 시점은 데이터 부족으로 오차가 컸지만 시점이 지날수록 실제 확산 형태를 근접하게 재현**
- **팀 노트북을 설정 파일 기반 Python 패키지로 재구성하고 테스트를 추가해 GitHub에 공개**

탄성파 해석 · GPlates · CCS · 중력 데이터 처리 · Python · TensorFlow · U-LSTM · ArcGIS

부트캠프 — SSAFY 16기 · 포스코 청년 AI·BigData 아카데미 30기 2025.06 – 진행 중

AI 프로젝트 · 삼성 청년 SW·AI 아카데미(SSAFY) · 포스코인재창조원·POSTECH · LLM 파인튜닝 · 데이터 파이프라인 · 백엔드

SSAFY 16기 관통 프로젝트 — ThRead: LLM 댓글 분석 파이프라인 · AI 도서 서비스

2026.07 – 진행 중 · 백엔드 · 데이터 파이프라인 담당 상세 ↗

- 관통 5 · 댓글 분석 파이프라인: Selenium으로 토스증권 커뮤니티 댓글 20개를 실시간 수집 → LLM이 부적절 댓글 번호만 판단하고 코드가 원본에서 제거 → pandas·정규식 정제 → IQR로 길이 이상치 기준 계산 → LLM 증강(개수 불일치 시 재시도) → 요약까지 단계별로 DB 저장
- 관통 5: 30초 이상 걸리는 작업의 진행 상황을 StreamingHttpResponse(NDJSON)+스레드·큐로 실시간 표시, 회사 없음·댓글 없음·LLM 실패 등 예외 상황을 구분해 안내
- 관통 4 · AI 도서 서비스: Django 커스텀 User·인증(오픈 리다이렉트 방어, HTTP Method 제한), 도서 CRUD, 위키피디아+LLM 작가 정보 생성, AI 도서 분석, 등록 도서 안에서만 추천하도록 검증한 AI 맞춤 추천, 알라딘 API 수집 명령
- 관통 4: Django TestCase와 mock으로 외부 API 없이 요구사항별 동작을 검증하는 자동 테스트 45개 작성
- 관통 1~3: JSON 도서 데이터 집계·정렬, 알라딘 API 수집(페이지네이션·중복 제거)+LLM 분류+gTTS 음성 변환, Bootstrap 반응형 웹페이지
- SSAFY 16기 AI 챌린지 (2026.09, 이미지+질문 4지선다 VQA, 팀 B022, 모델 학습·정제 담당): Qwen 27B/4B QLoRA·Gemma 31B 파인튜닝·OCR·멀티뷰(해상도·crop) 확률 앙상블로 Public Score 0.96961, GPU 없이 제출을 그대로 재현하는 스크립트(SHA-256 검증)까지 정리
- AI 코어 프로젝트 2 · 국보 입문 가이드 (2026.10, 개인): 국가유산청 오픈API(XML·EUC-KR)로 국보 목록·상세·공식 설명을 수집하고 국보명을 위키백과 문서에 자동 연결(일반명사 문서 오연결 방지) → ChromaDB 2개 컬렉션(국보 1건=청크 1개 / 위키·공식 설명 청크, ID 고정으로 재실행해도 중복 없음) → 질문마다 LLM이 검색어 추출·문서 선별 후 수집·색인(캐시·위키 장애 시 저장본 대체) → LangGraph retrieve→generate, 답변의 [근거 N]을 코드로 검사해 없는 번호 제거. 관람 해설·통설 판정·조건 탐색 기능, Streamlit+FastAPI 화면, 가짜 LLM·AppTest 기반 테스트 50개
- 새내기 캠프 3인 팀 프로젝트 「요기유(YOGIU)」 — 대전·충청 AI 여행 커뮤니티 웹 서비스, 개발·발표 담당: 카카오맵 API 기반 동선 플래너(찜·지도), OpenAI API 기반 AI 추천 여행 경로·꿈돌이 테마 챗봇, 테마별 관광 정보·축제·익명 게시판(댓글) 구현
- SSAFY AI 과정: Perceptron 구현, Transformer·BERT, RAG 검색, 프롬프트 엔지니어링, CNN·ViT / 일타싸피(삼각함수 기반 당구 AI), 배틀싸피 자동 플레이 봇(매 턴 재계획 BFS·탄약 판단·보급 암호 해독), SWEA 알고리즘
- **LLM 응답을 그대로 믿지 않고 코드로 검증하는 구조(번호 범위 검사, 뒤에서부터 제거, 등록 도서 id만 허용)로 안정적인 AI 기능 구현**
- **자동 테스트 관통 4 45개 · 관통 5 41개 통과, 관통 5 요구사항(F601~F611) 전부 구현 + 진행 상황 스트리밍·분석 이력·11가지 화면 테마 추가**

POSPEAK(KILL KONG) — 콩글리시 → 원어민 표현 교정 AI (FLAN-T5·Qwen + RAG)

2025.06 – 2025.08 · 데이터 구축·모델 실험 담당 상세 ↗

- [담당] 콩글리시·어색한 표현 → 원어민 표현 쌍으로 이루어진 단어 단위·문맥 단위 교정 데이터셋 구축 및 정제
- [담당] FLAN-T5를 LoRA로 파인튜닝하고, 학습 전략 3가지(단어→문맥 순차, 문맥→단어 순차, 통합 학습)를 BERTScore로 비교 실험
- [최종 서비스 'KILL KONG' — POSPEAK의 최종 버전] Qwen-2.5(1.5B) 파인튜닝(Friends 대사 기반 회화체 데이터 2,000건) + 콩글리시 RAG DB 630건 + 사용자별 개인화 메모리(오답노트)를 LangChain으로 결합, React Native·FastAPI 모바일 앱으로 구현
- [팀] 다국어 문장 임베딩(paraphrase-multilingual-MiniLM)과 FAISS로 유사 교정 예문 top-5를 검색해 모델 입력에 제공하는 RAG 파이프라인과 CLI 구성
- **순차 학습에서 앞 단계 성능이 떨어지는 catastrophic forgetting(BERTScore F1 0.47 → 0.32)을 실험으로 확인하고 통합 학습으로 전환**
- **통합 학습 모델의 검증 BERTScore F1 약 0.74 (순차 학습 0.32~0.47 대비 개선)**
- **팀 프로젝트 장려상 수상**

Python · LLM 파인튜닝 · QLoRA · RAG · FAISS · Django · Selenium · Pandas

PUBLICATIONS · 논문 및 발표

[논문] Wavenumber Correlation of Gravity and Magnetic Anomalies for Constraining Salt/Basement Structure in Offshore Abu Dhabi, *SPE Journal (SCIE) · 제1저자 · 게재 승인* 2026

[논문] 판구조론적 역사를 고려한 해외 대규모 이산화탄소 지중저장소 탄성파 해석 결과 분석, *지구물리와 물리탐사 27(1), 1–22 · 제1저자* 2024

[논문] 호주 깁스랜드 분지 CarbonNet CCS 프로젝트의 지질학적 분석, *지구물리와 물리탐사 · 제2저자* 2023

[발표] 인공지능과 디지털 트윈 융합을 통한 스마트 이산화탄소 저장소 운영 혁신 글로벌 인력양성사업 2024

[발표] 해외 CCS 프로젝트 실증 현황 및 국제공동연구 추진 전략 2024

[발표] 스마트 이산화탄소 지중저장 부지 운영을 위한 글로벌 인력양성사업 2024

[발표] 노르웨이 슬라이프너 지구물리 모니터링 적용사례 분석 및 국제공동연구 계획 2023

[발표] 사우디아라비아 딥러닝 기반 탄성파 영상화 연구 현황과 국제공동연구 계획 2023

[발표] GPU를 이용한 4차원 탄성파 모니터링 기술 적용성 검토 2022

SKILLS

언어·프레임워크 Python, TensorFlow, PyTorch, Hugging Face Transformers, Pandas, Django, Selenium, SQL

AI·데이터 분석 시공간 딥러닝(U-LSTM), CNN·LSTM, LLM 파인튜닝(FLAN-T5·Qwen·Gemma, LoRA/QLoRA), 멀티모델 확률 앙상블, RAG(LangGraph·ChromaDB·FAISS), 모델 평가(F1·BERTScore)

공간·지구물리 데이터 ArcGIS, 래스터 데이터 구축, 중력·자력 자료 처리, 파수 상관 분석, 탄성파 자료 해석, GPlates

도메인 CCS(이산화탄소 지중저장), 에너지·자원, 재난(산불) 예측

EDUCATION

삼성 청년 SW·AI 아카데미 (SSAFY) 2026.07 – 진행 중
16기 교육 과정

전북대학교 대학원 2024.03 – 2026.08
환경에너지융합학과 석사

포스코 청년 AI·BigData 아카데미 2025.06 – 2025.08
30기 수료

전북대학교 ~ 2024.02
자원에너지공학과 학사

AWARDS

포스코 청년 AI·BigData 아카데미 30기 프로젝트 장려상 2025.08

2024 한국지구물리·물리탐사학회(KSEG) 추계학술대회 학생 우수논문(포스터) 발표상 2024.11

CERTIFICATIONS

ADsP (데이터분석 준전문가) 2022.09