



서귀포지원 법원 AI Agent 개발 계획서



1. 데이터셋 구성표



공개 데이터셋 (Public Dataset)

분류	데이터 종류	출처	규모	형태	활용 목적
법령·판례	대한민국 법령집	국가법령정보센터	50만건+	JSON/XML	법률 조문 해석
	대법원 판례집	대법원 종합법률정보	100만건+	JSON/PDF	판례 분석 및 인용
	하급심 판결문	각급법원 판결서	500만건+	JSON/PDF	사건 유형별 판단 기준
	행정심판 결정례	국민권익위원회	10만건+	JSON	행정사건 참조
법원 업무	민사소송 절차 매뉴얼	법원행정처	1,000페이지	PDF/DOC	절차 안내 자동화
	가사사건 처리 지침	법원행정처	800페이지	PDF/DOC	가사사건 상담
	형사소송 절차 안내	법원행정처	1,200페이지	PDF/DOC	형사절차 설명
	등기사무 처리 규정	법원행정처	600페이지	PDF/DOC	등기업무 자동화
민원 안내	법원 민원 FAQ	대법원 홈페이지	5,000건	HTML/JSON	기본 질의응답
	소송비용 계산표	법원행정처	100페이지	Excel/JSON	비용 계산 자동화
	서식 모음집	각급법원	1,000건	HWP/PDF	서식 제공 및 작성
지역 정보	제주특별법 조문	제주도청	500조문	JSON	제주 특례법 적용

언어 모델	제주도 조례집	제주도의회	1,000건	JSON/PDF	지역 법규 해석
	법률 용어사전	법제처	50,000어	JSON	전문용어 이해
	한국어 법률 코퍼스	AI Hub	1TB	Text	법률 언어 모델링

🔒 비공개 데이터셋 (Private Dataset)

분류	데이터 종류	출처	규모	형태	활용 목적	보안등급
사건 처리	서귀포지원 사건 접수 현황	서귀포 지원	100,000건	DB	사건 분류 및 배당	대외비
	과거 판결문 (익명화)	서귀포 지원	50,000건	JSON	지역 특성 반영 판단	대외비
	조정·화해 사례	서귀포 지원	20,000건	DB	분쟁 해결 솔루션	대외비
내부 업무	사무분장 및 담당자 정보	서귀포 지원	500건	DB	업무 배정 및 연계	대외비
	내부 업무 매뉴얼	서귀포 지원	2,000페이지	DOC	내부 절차 자동화	대외비
	사건 진행 현황 데이터	서귀포 지원	실시간	DB	실시간 진행 상황	대외비
민원 이력	과거 민원 상담 기록	서귀포 지원	30,000건	DB	상담 품질 향상	대외비
	전화 상담 음성 데이터	서귀포 지원	10,000시간	WAV	음성 인식 및 응답	대외비
	온라인 채팅 로그	서귀포 지원	50,000건	JSON	챗봇 대화 개선	대외비
지역 특성	제주도 특수 사건 유형	서귀포 지원	5,000건	DB	제주 특화 업무	대외비

7일 차	공개 데이터 전처리 II	• FAQ 데이터 구조화 • 절차 매뉴얼 토큰화 • 서식 정보 추출 • 용어사전 매핑	데이터 엔지니어	구조화된 매뉴얼
8일 차	비공개 데이터 수집	• 서귀포지원 DB 연결 • 사건 데이터 추출 • 개인정보 마스킹 • 데이터 검증	시스템 관리자	익명화된 사건 DB
9일 차	비공개 데이터 전처리 I	• 사건 분류 라벨링 • 판결문 요약 생성 • 진행 단계 태깅 • 시간 정보 정규화	법무팀 + AI팀	라벨링된 사건 데이터
10 일 차	비공개 데이터 전처리 II	• 민원 상담 분류 • 질의-응답 쌍 생성 • 대화 컨텍스트 추출 • 감정 분석 태깅	법무팀 + AI팀	Q&A 데이터셋
11 일 차	제주 특화 데이터 처리	• 제주특별법 매핑 • 관광 분쟁 패턴 분석 • 부동산 특성 태깅 • 지역 조례 연결	제주 법무팀	제주 특화 DB
12 일 차	학습 데이터 셋 구성	• Train/Validation/Test 분할 • 데이터 밸런싱 • 증강 데이터 생성 • 최종 품질 검증	데이터 사이언 티스트	학습용 데이터셋
13 일 차	프롬프 트 엔지니어 링	• 법률 상담용 프롬프트 설계 • 절차 안내용 템플릿 작성 • 서식 작성 가이드 개발 • Few-shot 예제 준비	AI 연구원	프롬프트 라이브러 리
14 일 차	평가 데이터 셋 구축	• 법률 지식 평가 세트 • 절차 안내 평가 세트 • 민원 응답 평가 세트 • 제주 특화 평가 세트	평가팀	평가 벤치마크
15 일 차	데이터 최종 검증	• 전체 데이터 통합 • 일관성 검사 • 누락 데이터 보완 • 학습 준비 완료	전체팀	최종 데이터셋

🎯 Fine-tuning 실행 단계 (16-25일차)

일 차	작업 내용	세부 사항	담당 자	결과물
--------	-------	-------	---------	-----

16 일차	LoRA Fine-tuning 시작	• LoRA 설정 최적화 • 첫 번째 학습 실행 • 학습 곡선 모니터링 • 초기 성능 평가	AI 엔지 니어	1차 모델 체크포인트
17 일차	법률 지식 특화 학습	• 판례 학습 집중 • 법령 해석 강화 • 용어 정의 학습 • 중간 평가 실시	AI 엔지 니어	법률 지식 모델
18 일차	절차 안내 특화 학습	• 민사/형사/가사 절차 • 등기 업무 절차 • 서식 작성 가이드 • 비용 계산 학습	AI 엔지 니어	절차 안내 모델
19 일차	민원 상담 특화 학습	• 대화형 응답 학습 • 감정 대응 훈련 • 다단계 질의 처리 • 상황별 응답 최적화	AI 엔지 니어	상담 특화 모델
20 일차	제주 특화 학습	• 제주특별법 적용 • 관광 분쟁 해결 • 부동산 특성 반영 • 지역 조례 연동	AI 엔지 니어	제주 특화 모델
21 일차	다중 태스크 통합 학습	• 전체 기능 통합 • 태스크 간 연결 강화 • 컨텍스트 이해 향상 • 일관성 유지 훈련	AI 엔지 니어	통합 모델
22 일차	하이퍼파라 미터 튜닝	• Learning Rate 최적화 • Batch Size 조정 • Dropout 설정 • 정규화 강화	AI 엔지 니어	최적화된 설정
23 일차	추가 데이터 학습	• 누락된 케이스 보완 • 엡지 케이스 학습 • 오류 케이스 수정 • 성능 미달 영역 강화	AI 엔지 니어	보완 학습 완료
24 일차	모델 검증 및 디버깅	• 전체 성능 평가 • 오답 케이스 분석 • 편향성 검사 • 안전성 검증	평가 팀	검증 리포트
25 일차	최종 모델 생성	• 최고 성능 체크포인트 선택 • 모델 압축 및 최적화 • 배포용 형태 변환 • 백업 모델 생성	AI 엔지 니어	최종 Fine-tuned 모델

 테스트 및 배포 준비 (26-30일차)

일차	작업 내용	세부 사항	담당 자	결과물
26일 차	통합 테스트	• 실제 민원 시나리오 테스트 • 다양한 질의 유형 검증 • 응답 품질 평가 • 속도 성능 측정	QA팀	테스트 결과서
27일 차	법무팀 검증	• 법률 정확성 검토 • 절차 안내 정확성 확인 • 제주 특화 내용 검증 • 개선 사항 도출	법무 팀	법무 검증서
28일 차	사용자 테스트	• 서귀포지원 직원 테스트 • 실제 민원인 파일럿 • 사용성 평가 • 피드백 수집	사용 자팀	사용자 피드백
29일 차	최종 개선	• 테스트 결과 반영 • 마이너 버그 수정 • 성능 최적화 • 문서화 완료	전체 팀	배포 준비 완료
30일 차	배포 및 운영 준비	• 프로덕션 환경 배포 • 모니터링 시스템 구축 • 운영 매뉴얼 작성 • 교육 자료 준비	운영 팀	서비스 런칭

성능 목표 및 평가 지표

목표 성능

- 법률 지식 정확도: 95% 이상
- 절차 안내 완성도: 98% 이상
- 민원 응답 만족도: 90% 이상
- 응답 속도: 3초 이내
- 제주 특화 대응: 85% 이상

평가 메트릭

- **BLEU Score:** 번역 품질 평가
- **ROUGE Score:** 요약 품질 평가
- **BERTScore:** 의미적 유사도 평가
- **Human Evaluation:** 전문가 평가
- **User Satisfaction:** 사용자 만족도

예상 비용 및 자원

하드웨어

- GPU 서버: A100 40GB × 4 (월 임대료 약 800만원)
- 스토리지: 100TB SSD (500만원)
- 네트워크: 전용 회선 (월 200만원)

인력

- AI 엔지니어 2명 × 30일
- 데이터 엔지니어 2명 × 15일
- 법무 전문가 1명 × 20일
- 프로젝트 매니저 1명 × 30일

기대 효과

- 업무 효율성: 70% 향상
- 민원 처리 시간: 50% 단축
- 인력 절감: 30% 감소
- 시민 만족도: 40% 향상