



Travel AI

검색부터 맞춤 견적 생성까지, 막힘없이 이어지는 AI 여행 상담 챗봇

01. 서비스 개요 및 기획 배경

팀원소개



김유신

ERD 설계
RAG 구축
백엔드



임규리

AI 모델링
RAG 구축
프론트엔드

Table of Contents

01. 서비스 개요 및 기획 배경

02. 프로젝트 관리

03. 시스템 아키텍처 및
DB 설계

04. AI 아키텍처 진화 및
핵심 파이프라인

05. 시연

06. 고도화

01. 서비스 개요 및 기획 배경

여행업의 현실과 문제점 (Pain Points)

"왜 여행사는 여전히 바쁘고 비효율적인가?"

디지털 전환이 가장 더딘 Golf & Luxury Travel 시장의 3대 문제점에 주목.



1 데이터 표준의 부재 (Lack of Data Standardization)

- 현황:
랜드사마다 제각각인 견적 양식(PDF, 엑셀 등) 혼재.
- 문제:
수기 입력 의존으로 전체 리소스 40% 낭비
- 결과:
데이터 정합성 결여 및 휴먼 에러 발생 상존.

2 수동 검색의 비효율 (Inefficiency of Manual Search)

- 현황:
고객 질문 시 상담원이 엑셀/웹을 일일이 뒤져야 함.
- 문제:
상품이 많을수록 검색 시간이 과다 소요되어 실시간 응대 불가.
- 결과:
고객 대기 시간 증가로 인한 이탈율(Bounce Rate) 상승.

3 응대 확장성의 한계 (Limit of Scalability)

- 현황:
숙련된 상담원만이 1:1로 맞춤 견적(PDF) 작성 가능.
- 문제:
고객이 늘어나면 상담원도 같이 늘려야 하는 고비용 구조.
- 결과:
업무 외 시간에 들어오는 수요 대응 불가 및 매출 기회 손실.

01. 서비스 개요 및 기획 배경

해결 솔루션: As-Is vs To-Be

Hybrid AI 기술을 통해 기존의 비효율적 수동 프로세스를 완전 자동화된 데이터 파이프라인으로 재설계

비교 항목	AS-IS: 기존 여행사 ERP / 예약 시스템	TO-BE: AI Travel Agent
1. 검색 방식	<p>[키워드 의존 & 결과 제한]</p> <ol style="list-style-type: none">사용자가 정확한 상품명이나 지역명(예: "다낭 A골프장")을 모르면 검색 불가."부모님 효도 여행" 같은 추상적 니즈는 처리 못하고 '결과 없음' 반환.오타나 띄어쓰기 오류 시 검색 실패 (Rigid Rule-based).	<p>[맥락 기반 추론]</p> <p>불완전한 발화의 의도를 파악하는 시맨틱 검색 고객의 숨은 의도 파악 및 추천.</p>
2. 정보 입력	<p>[복잡한 텁스(Depth) & 수동 입력]</p> <ol style="list-style-type: none">날짜, 인원, 국가, 호텔 등 10개 이상의 필터를 일일이 클릭/입력해야 함.조건 변경 시 처음부터 다시 세팅해야 하는 번거로움.상담원이 고객과 통화하며 시스템에 받아 적는 Double Work 발생.	<p>[자동 슬롯 필링]</p> <p>자연어 대화에서 예약 정보를 자동 추출/입력. Zero-UI (복잡한 클릭 없음).</p>
3. 프로세스	<p>[단절된 데이터 & 수동 매칭]</p> <ol style="list-style-type: none">항공, 숙박, 골프 예약 화면이 분리되어 있어 교차 확인(Cross-check) 필요.예약 불가능 시 대안을 상담원이 직접 머리 써서 찾아야 함.상담원 개인 역량에 따른 서비스 품질 편차 심함.	<p>[One-Stop Automation]</p> <p>Task Locking으로 예약 완결까지 자동 유도. 데이터 통합으로 즉각적인 대안 제시.</p>

01. 서비스 개요 및 기획 배경

서비스 개요 (Service Overview)

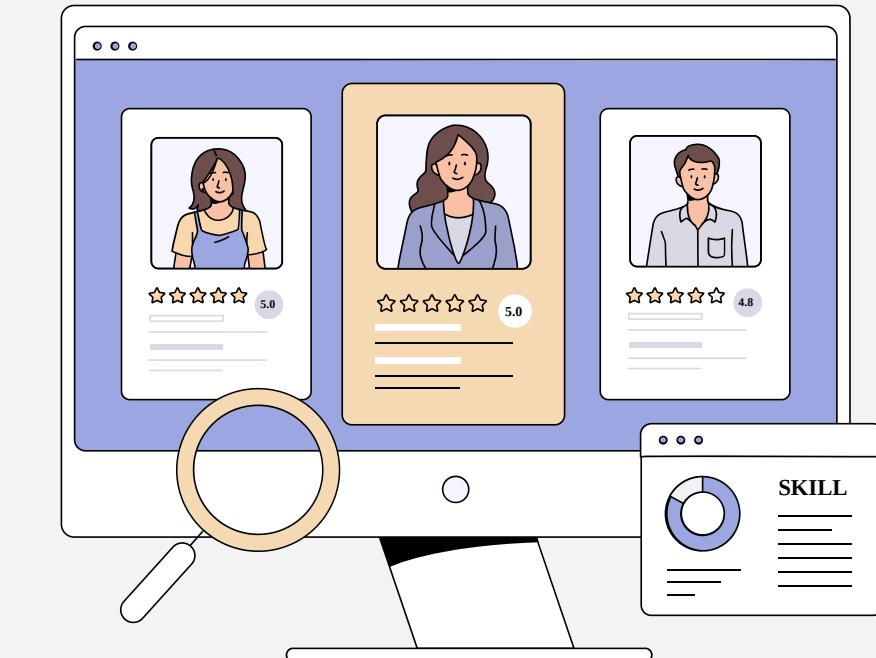
랜드사의 비정형 데이터를 자산화하고, 고객 상담부터
예약까지의 전 과정을 자동화하는 지능형 여행 업무 플랫폼

표준 포맷 도입을 통한 등록 자동화

- 표준 JSON 파이프라인 도입으로 수기 입력 전면 대체
- Schema 매팅 엔진을 통한 즉시 파싱 및 DB 적재 자동화

서비스 내 완결형 채팅 환경

- 휘발성 상담 데이터를 내재화(In-House)하여 AI 자산으로 전환
- Context(화면)와 대화를 결합해 견적서 즉시 생성
- 대화가 곧바로 예약 데이터로 변환되는 Actionable Chat 구현



대기 없는 실시간 AI 상담

- RAG+LLM 기반 실시간 응대 및 맞춤형 상품 추천
- 대기 없는 실시간 데이터 검색으로 신속·정확한 정보 제공

02. 프로젝트 관리

요구사항 정의서

요구사항 정의서				
프로젝트 명: Travel_AI 문서업데이트 : 2025.12.17 작성자 : 임규리				
No	구분	요구사항 ID	요구사항 명	요구사항 상세 설명
1	기능	RQ-01	프로젝트 개요	프로젝트명: Travel_AI (지능형 여행 업무 자동화 플랫폼) 개발 환경: - BE: Python (FastAPI, Flask), Kafka, WebSocket, SQLAlchemy - FE: React, TailwindCSS, JavaScript - AI: PyTorch (KoELECTRA/KoBART), Gemini API, LangChain, PaddleOCR/EasyOCR, BGE-3 - DB: MySQL, Vector DB (Chroma/Pinecone) 목적: 골프/낚시 여행 시장의 비효율(리소스 40% 낭비)을 해소하고, Kafka/WebSocket 기반의 실시간 상담 처리 및 예약 데이터 자산화를 구현함.
2			시스템 아키텍처	1. MSA 지향 아키텍처: FastAPI를 메인 게이트웨이로 두고, 실시간 제팅(SocketIO)과 AI 모델 서빙을 분리하여 안정성을 확보한다. 2. 데이터 파이프라인: Kafka를 도입하여 대량의 상담 로그와 예약 데이터를 실시간으로 스트리밍 및 적재하는 구조를 갖춘다.
3		RQ-03	상품 등록 및 데이터 구조화	1. 표준 JSON 파이프라인 - 관리자가 입력한 데이터를 시스템 표준 JSON 스키마로 매핑하여 DB 적재를 자동화한다. 2. 비정형 데이터 처리 기반 - PaddleOCR, pdfminer 등 파싱 라이브러리를 아키텍처에 포함하여, 형후 PDF/이미지 자동 인식을 위한 확장 환경(Foundation)을 제공한다.
4			의도 분류 및 라우팅	1. 고속 분류 모델: KoELECTRA 기반의 분류 모델을 통해 사용자 발화를 0.1초 내에 분석한다. 2. 트렉 라우팅: 분류 결과에 따라 Type A(예약), Type B(추천), Type C(건적), 상담원 연결(Human-Handoff) 트랙으로 자동 분기한다.
5			실시간 AI 상담 (RAG)	1. 벡터 검색: LangChain 및 Vector DB를 연동하여, 사용자 질문에 대해 여행 상품 정보를 실시간으로 검색 (Retrieval)하여 답변한다. 2. 상품 추천: 고객의 모호한 요구사항을 벡터 유사도 기반으로 분석하여 최적의 골프장/호텔을 추천한다.
6			자동 건적 생성 (LLM)	1. 통합 LLM 파이프라인: Gemini와 LangChain을 활용, 긴 상담 문맥(Context)을 분석하여 여행 일정표가 포함된 건적서를 자동 생성한다. 2. 출력 제어: JSON Mode 및 풋볼프트 언저리닝을 통해 할루시네이션을 억제하고 정형화된 건적 데이터를 출력한다.
7			예약 정보 추출 (NER)	1. 슬롯 팔링: KoELECTRA NER 모델을 사용하여 대화 중 [날짜], [장소], [인원] 정보를 정밀하게 추출한다. 2. 정규화: 추출된 언티티를 시스템 포맷(YYYY-MM-DD 등)으로 변환하여 예약 DB에 매핑한다.
8			Actionable Chat	1. 실시간 통신: WebSocket 및 Flask-SocketIO를 통해 끊김 없는 실시간 양방향 상담 환경을 제공한다.
9		RQ-04	성능 및 안정성	1. 비동기 처리: AsyncIO 및 FastAPI의 비동기 기능을 활용하여 다수의 동시 접속 상담 트래픽을 지원 없이 처리한다. 2. 내결함성: Kafka를 통해 트래픽 폭주 시에도 데이터 유실 없이 메시지를 버퍼링(Buffering)한다.
10			보안 및 운영	1. 데이터 암호화: 개인정보(전화번호 등)는 암호화 모듈(cryptography)을 통해 DB에 저장한다. 2. 모니터링: 시스템 리소스 및 AI 토큰 사용량을 모니터링할 수 있는 대시보드(Admin)를 제공한다.

기능 정의서(SW)		
프로젝트 명: TRVEL_AI 문서업데이트 : 2025.12.11		
구분	메뉴	기능 설명
User (Client)	채팅 인터페이스	<ul style="list-style-type: none">사용자 자연어 텍스트 입력 및 전송AI 답변 실시간 출력 및 UI 표시
AI Core	의도 분류 (Router)	<ul style="list-style-type: none">사용자 입력 텍스트 분석하여 6가지 의도로 자동 분류 (Greeting, Complaint, Booking, Recommend, Custom, Irrelevant)현재 진행 중인 태스크(Task Lock) 유무 체크
AI Core	슬롯 팔링 (Engine)	<ul style="list-style-type: none">대화 문맥에서 필수 정보(출발지, 인원, 날짜, 기간 등) 자동 추출누락된 필수 정보 식별 및 역질문 생성JSON 포맷으로 정형 데이터 변환 및 간결화
AI Core	기억 관리 (Memory)	<ul style="list-style-type: none">세션(Session)별 대화 히스토리 저장 및 관리지시 대명사("거기", "그것") 및 이전 문맥 추론 기능 제공
Function	상품 추천 (RAG)	<ul style="list-style-type: none">사용자 요청(자연어)에 기반한 벡터 유사도 검색 수행검색된 여행 상품 정보 요약 및 추천 사유 생성[Fallback] 검색 결과 없을 시 자동으로 '맞춤 견적' 모드로 전환
Function	간편 예약 (Booking)	<ul style="list-style-type: none">구체적 상품 지정 시 예약 모드로 진입 (Task Locking)예약 필수 정보(출발지, 인원) 수집 완료 시 예약 접수 처리추가 요청사항(숙소, 식사 등) 확인 질문 자동 생성
Function	맞춤 견적 (Custom)	<ul style="list-style-type: none">추천 거절 또는 맞춤 요청 시 견적 모드로 진입 (Task Locking)5대 필수 항목(출발지, 인원, 국가, 기간, 출발일) 완결 수집특수 요청사항(휠체어, 알러지 등) 메모리 저장
Function	고객 응대 (CS)	<ul style="list-style-type: none">[Greeting] 여행 관련 오프닝 인사 및 서비스 안내[Complaint] 불만 키워드 감지 시 위로 메시지 출력 및 상담원 연결 프로세스 안내
System	예외 처리	<ul style="list-style-type: none">여행 무관 주제(Irrelevant) 입력 시 답변 거절 처리정보 인식 실패 시 재질문 유도 (Error Handling)

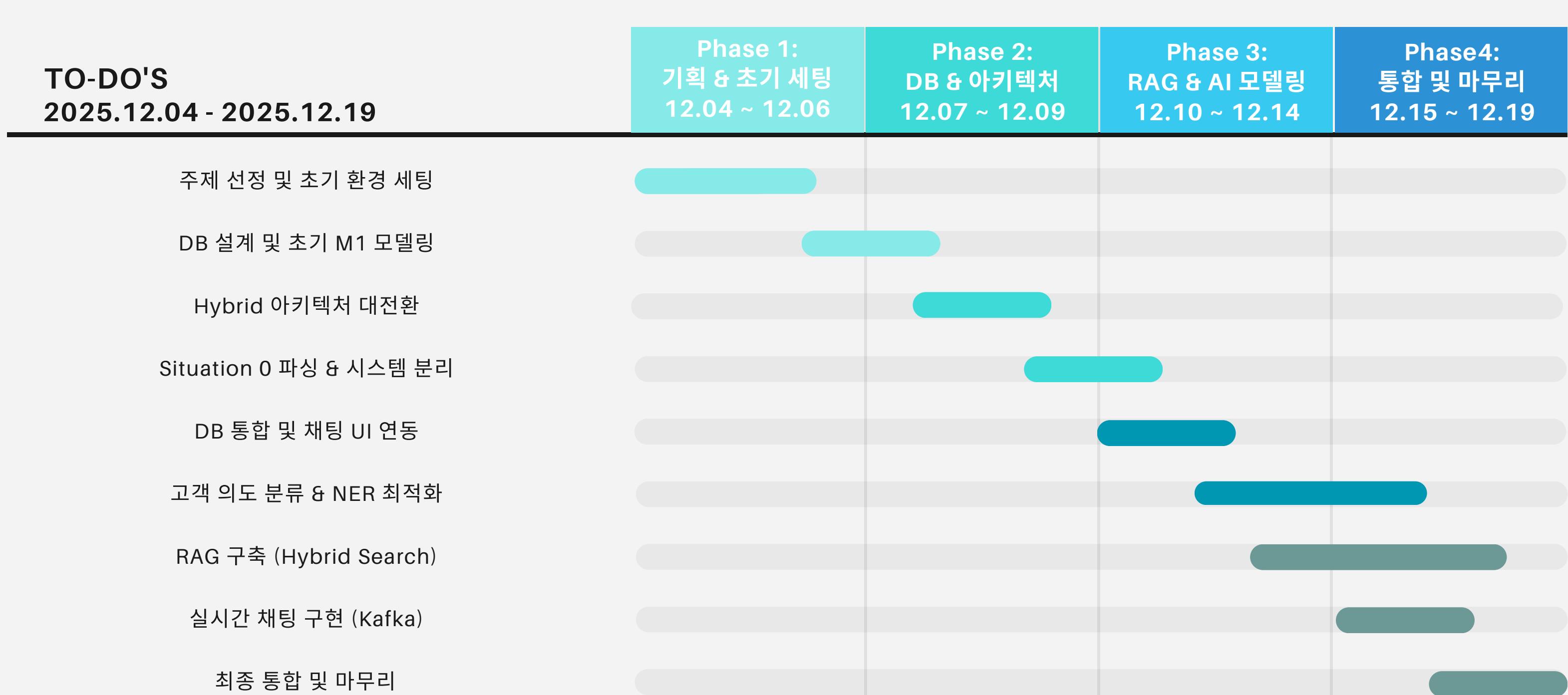
02. 프로젝트 관리

WBS

2	설계			2025-12-04	2025-12-15	12		0.00%		0.00%
2.1	시스템 아키텍처 설계			2025-12-04	2025-12-09	6	6	100.00%		100.00%
2.1.1	초기 모델 아키텍처 선정 (Transformer/NLP 중심)			2025-12-04	2025-12-04	1	1	100.00%	임규리	100.00%
2.1.2	개발 환경 및 Docker 컨테이너 구조 정의			2025-12-04	2025-12-04	1	1	100.00%	임규리	100.00%
2.1.3	Hybrid AI 아키텍처 재설계 (Local + Gemini API)			2025-12-08	2025-12-08	1	1	100.00%	김유신	100.00%
2.1.3	서비스 이원화 설계 (Client Web / Admin ERP 포트 분리)			2025-12-09	2025-12-09	1	1	100.00%	김유신	100.00%
2.3	DB 설계			2025-12-05	2025-12-15	11	11	100.00%		0.00%
2.3.1	상품/건적 데이터 구조 정의 (JSON Key Mapping)			2025-12-05	2025-12-05	1	1	100.00%	김유신	100.00%
2.3.2	여행사-랜드사-고객 데이터 연관 관계(ERD) 설계			2025-12-05	2025-12-05	1	1	100.00%	김유신	100.00%
2.3.3	채팅 시스템 DB 분리 설계 (Rooms/Logs 테이블 정규화)			2025-12-14	2025-12-14	1	1	100.00%	김유신	0.00%
2.3.4	보안 테이블 설계 (여권 AES-256 암호화 및 PW 해싱)			2025-12-15	2025-12-15	1	1	100.00%	김유신	0.00%
2.4	AI/로직 설계			2025-12-08	2025-12-15	8	8	100.00%		100.00%
2.4.1	업무 상황별(Situation 0~2) AI 처리 프로세스 정의			2025-12-08	2025-12-08	1	1	100.00%	임규리	100.00%
2.4.2	비정형 데이터 파싱 전략 수립 (Text/Vision 분기 로직)			2025-12-08	2025-12-09	2	8	400.00%	김유신	100.00%
2.4.3	RAG 하이브리드 검색 로직 설계 (Vector + Re-ranking)			2025-12-13	2025-12-13	1	1	100.00%	김유신	100.00%
2.4.4	LLM 라우팅 및 State Machine(상태 유지) 로직 설계			2025-12-15	2025-12-15	1	1	100.00%	임규리	100.00%
2.5	웹 디자인 설계			2025-12-04	2025-12-05	1	1	100.00%		100.00%
2.5.1	화면 구성 및 정보 구조			2025-12-04	2025-12-04	-	1	#DIV/0!	임규리	100.00%
2.5.2	와이어프레임 설계			2025-12-04	2025-12-04	-	1	#DIV/0!	임규리	100.00%
2.5.3	UX 흐름 설계			2025-12-04	2025-12-05	2	2	100.00%	임규리	100.00%
2.5.4	UI 컴포넌트 설계			2025-12-04	2025-12-05	2	2	100.00%	임규리	100.00%

02. 프로젝트 관리

간트차트



03. 시스템 아키텍처 및 데이터베이스 설계

기술 스택

Frontend



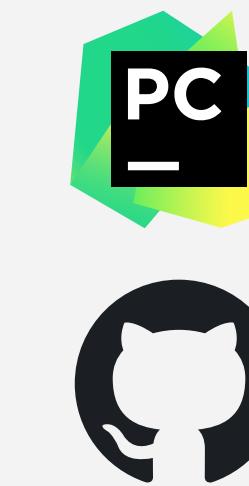
Backend



AI & Data



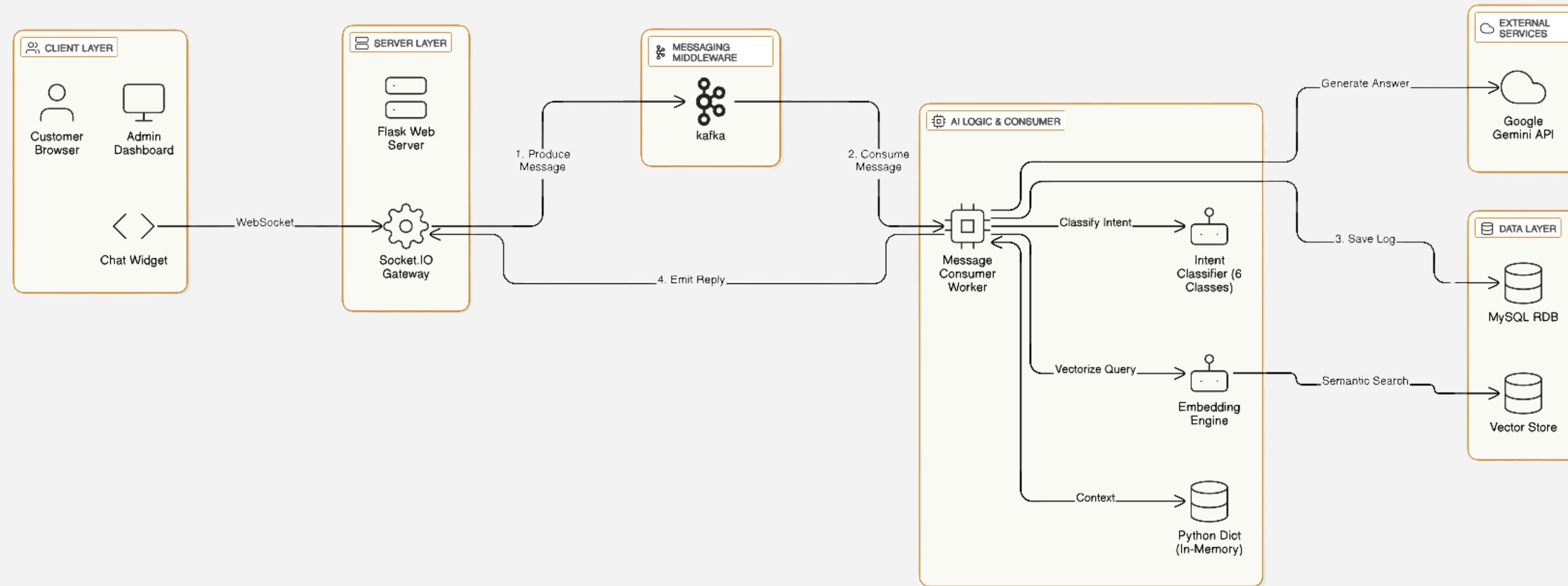
IDE & 협업도구



03. 시스템 아키텍처 및 데이터베이스 설계

시스템 인프라 및 데이터 구조

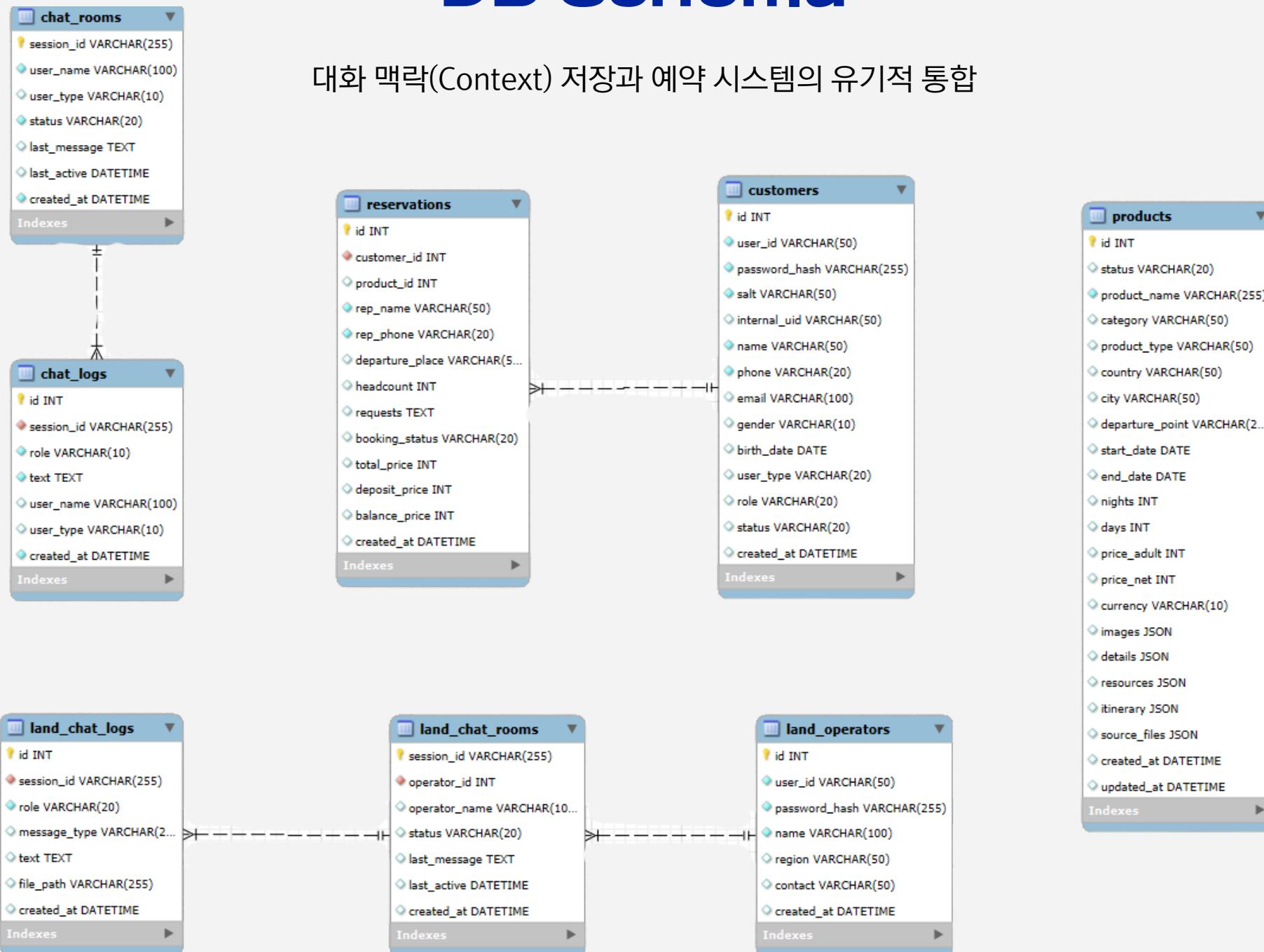
실시간 데이터 처리와 효율성을 고려한 시스템 아키텍처



03. 시스템 아키텍처 및 데이터베이스 설계

DB Schema

대화 맥락(Context) 저장과 예약 시스템의 유기적 통합



03. 시스템 아키텍처 및 데이터베이스 설계

비정형 데이터의 정형화(JSON)

OCR → Human-in-the-Loop 파이프라인 구축

 <p>서울:(04043) 서울특별시 마포구 전다리로 33, 3층 부산:(48055) 부산시 해운대구 해운대로 161번길 19 센텀해운대 402호 www.teratour.co.kr Tel. 02-6956-1451 Fax. 02-6008-1451</p>																					
<p>[25년10월~26년03월] 겨울인기NO1. 휴양과골프 오키나와시내3색</p> <p>골프라운딩후 온천여행, 시내관광까지 즐길수 있는 오감만족 오키나와</p>																					
<p>포함 사항 3성급, 4성급, 5성급 호텔(조식포함), 그린피+카트피+악카피, 숭영차량(일본인기사)</p> <p>불포함 사항 항공권, 공항세 및 유류할증료, 캐피, NO가이드, 여행자보험</p>																					
<p>비고 1. 3성급 호텔은 산스이 나하, 호텔 OR 숲비타 호텔 나하 기준입니다 나하시티 센터에 위치한 산스이호텔은 호텔에서 나하하시내까지 무료셔틀 운행합니다 (단 시내에서 → 호텔까지는 개인이동) 4성급 호텔은 휴잇로 오키나와 기준입니다 5성급 호텔은 플렉티 호텔입니다 (대체호텔 없음) 2인, 3인 플레이시 골프 할인 비용 지상비요금표에 종요금으로 표기되어있습니다 2인, 3인 숭영시 지상비 요금표에 종요금으로 표기되어있습니다 3. 숭영의 골프장은 오션캐슬 CC / 오키나와 컨트리 CC 또는 팔힐즈 GC / 쿠라오차드 GC 기준입니다 (골프장은 지정불가능, 시간지정불가능이며, 18홀스루플레이, 중간휴식없는 플레이 조건입니다) 4. 오전출발 항공출발일경우 인천항공, 부산출발 전에 이스타 항공은 첫날 18홀이 가능합니다 (인천출발 OZLU 항공시간대는 첫날 9홀만 가능합니다) 오후출발은 인천항공 제주항공 페탄으로 (0/18/18/18) 가능합니다</p>																					
<table border="1"> <thead> <tr> <th>날짜</th> <th>교통</th> <th>시간</th> <th>(온전출발 - 오전리턴 참고일정표)</th> <th>식사</th> </tr> </thead> <tbody> <tr> <td>1일차</td> <td>오전출발편 송영차량</td> <td></td> <td>부산 또는 인천 국제공항 출발 나하 국제공항 도착 일본 숭영차량 기사 미팅 후 골프장 이동 ▶ 오키나와컨트리, 오션캐슬, 팔힐즈, 쿠라오차드 CC 18홀(카트+클럽포함) 간단드레스, 호텔 이동, 호텔 체크인 후 자유식식 및 자유 시간 TEL: 3성급 OR 4성급 OR 5성급 (2인1실)</td> <td>중:불포함 석:불포함</td> </tr> <tr> <td>체류일 (2,3일차)</td> <td>송영차량</td> <td></td> <td>호텔 조식 / 기사미팅 / 골프장 이동 ▶ 오키나와컨트리, 오션캐슬, 팔힐즈, 쿠라오차드 CC 18홀(카트+클럽포함) 간단드레스, 호텔 이동, 호텔 체크인 후 자유식식 및 자유 시간 HOTEL: 3성급 OR 4성급 OR 5성급 (2인1실)</td> <td>조:호텔식 중:불포함 석:불포함</td> </tr> <tr> <td>귀국일</td> <td>송영차량 오전리턴편</td> <td></td> <td>호텔 조식 후 체크아웃 후 기사미팅 후 공항으로 이동 (약 10분소요) 나하국제공항 출발 부산 또는 인천 국제공항</td> <td>조:호텔식</td> </tr> </tbody> </table>		날짜	교통	시간	(온전출발 - 오전리턴 참고일정표)	식사	1일차	오전출발편 송영차량		부산 또는 인천 국제공항 출발 나하 국제공항 도착 일본 숭영차량 기사 미팅 후 골프장 이동 ▶ 오키나와컨트리, 오션캐슬, 팔힐즈, 쿠라오차드 CC 18홀(카트+클럽포함) 간단드레스, 호텔 이동, 호텔 체크인 후 자유식식 및 자유 시간 TEL: 3성급 OR 4성급 OR 5성급 (2인1실)	중:불포함 석:불포함	체류일 (2,3일차)	송영차량		호텔 조식 / 기사미팅 / 골프장 이동 ▶ 오키나와컨트리, 오션캐슬, 팔힐즈, 쿠라오차드 CC 18홀(카트+클럽포함) 간단드레스, 호텔 이동, 호텔 체크인 후 자유식식 및 자유 시간 HOTEL: 3성급 OR 4성급 OR 5성급 (2인1실)	조:호텔식 중:불포함 석:불포함	귀국일	송영차량 오전리턴편		호텔 조식 후 체크아웃 후 기사미팅 후 공항으로 이동 (약 10분소요) 나하국제공항 출발 부산 또는 인천 국제공항	조:호텔식
날짜	교통	시간	(온전출발 - 오전리턴 참고일정표)	식사																	
1일차	오전출발편 송영차량		부산 또는 인천 국제공항 출발 나하 국제공항 도착 일본 숭영차량 기사 미팅 후 골프장 이동 ▶ 오키나와컨트리, 오션캐슬, 팔힐즈, 쿠라오차드 CC 18홀(카트+클럽포함) 간단드레스, 호텔 이동, 호텔 체크인 후 자유식식 및 자유 시간 TEL: 3성급 OR 4성급 OR 5성급 (2인1실)	중:불포함 석:불포함																	
체류일 (2,3일차)	송영차량		호텔 조식 / 기사미팅 / 골프장 이동 ▶ 오키나와컨트리, 오션캐슬, 팔힐즈, 쿠라오차드 CC 18홀(카트+클럽포함) 간단드레스, 호텔 이동, 호텔 체크인 후 자유식식 및 자유 시간 HOTEL: 3성급 OR 4성급 OR 5성급 (2인1실)	조:호텔식 중:불포함 석:불포함																	
귀국일	송영차량 오전리턴편		호텔 조식 후 체크아웃 후 기사미팅 후 공항으로 이동 (약 10분소요) 나하국제공항 출발 부산 또는 인천 국제공항	조:호텔식																	
<table border="1"> <thead> <tr> <th>날짜</th> <th>교통</th> <th>시간</th> <th>(오후출발 - 오후리턴 참고일정표)</th> <th>식사</th> </tr> </thead> <tbody> <tr> <td>1일차</td> <td>오후출발편 송영차량</td> <td></td> <td>인천 국제공항 출발 나하 국제공항 도착 일본 숭영차량 기사 미팅 후 호텔이동 호텔하차 호텔 개별체크인 후 자유식식 및 자유 시간 HOTEL: 3성급 OR 4성급 OR 5성급 (2인1실)</td> <td>석:불포함</td> </tr> <tr> <td>체류일 (2,3일차)</td> <td>송영차량</td> <td></td> <td>호텔 조식 / 기사미팅 / 골프장 이동 ▶ 오키나와컨트리, 오션캐슬, 팔힐즈, 쿠라오차드 CC 18홀(카트+클럽포함) 간단드레스, 호텔 이동, 호텔 체크인 후 자유식식 및 자유 시간 HOTEL: 3성급 OR 4성급 OR 5성급 (2인1실)</td> <td>조:호텔식 중:불포함 석:불포함</td> </tr> <tr> <td>귀국일</td> <td>송영차량 오후리턴편</td> <td></td> <td>호텔 조식 후 체크아웃 후 기사미팅 공항으로 이동 나하국제공항 출발 인천 국제공항</td> <td>조:호텔식</td> </tr> </tbody> </table>		날짜	교통	시간	(오후출발 - 오후리턴 참고일정표)	식사	1일차	오후출발편 송영차량		인천 국제공항 출발 나하 국제공항 도착 일본 숭영차량 기사 미팅 후 호텔이동 호텔하차 호텔 개별체크인 후 자유식식 및 자유 시간 HOTEL: 3성급 OR 4성급 OR 5성급 (2인1실)	석:불포함	체류일 (2,3일차)	송영차량		호텔 조식 / 기사미팅 / 골프장 이동 ▶ 오키나와컨트리, 오션캐슬, 팔힐즈, 쿠라오차드 CC 18홀(카트+클럽포함) 간단드레스, 호텔 이동, 호텔 체크인 후 자유식식 및 자유 시간 HOTEL: 3성급 OR 4성급 OR 5성급 (2인1실)	조:호텔식 중:불포함 석:불포함	귀국일	송영차량 오후리턴편		호텔 조식 후 체크아웃 후 기사미팅 공항으로 이동 나하국제공항 출발 인천 국제공항	조:호텔식
날짜	교통	시간	(오후출발 - 오후리턴 참고일정표)	식사																	
1일차	오후출발편 송영차량		인천 국제공항 출발 나하 국제공항 도착 일본 숭영차량 기사 미팅 후 호텔이동 호텔하차 호텔 개별체크인 후 자유식식 및 자유 시간 HOTEL: 3성급 OR 4성급 OR 5성급 (2인1실)	석:불포함																	
체류일 (2,3일차)	송영차량		호텔 조식 / 기사미팅 / 골프장 이동 ▶ 오키나와컨트리, 오션캐슬, 팔힐즈, 쿠라오차드 CC 18홀(카트+클럽포함) 간단드레스, 호텔 이동, 호텔 체크인 후 자유식식 및 자유 시간 HOTEL: 3성급 OR 4성급 OR 5성급 (2인1실)	조:호텔식 중:불포함 석:불포함																	
귀국일	송영차량 오후리턴편		호텔 조식 후 체크아웃 후 기사미팅 공항으로 이동 나하국제공항 출발 인천 국제공항	조:호텔식																	
<p>*상기 일정은 항공 및 현지사정에 의해 변동될 수 있습니다.</p>																					

 <p>서울:(04043) 서울특별시 마포구 전다리로 33, 3층 부산:(48055) 부산시 해운대구 해운대로 161번길 19 센텀해운대 402호 www.teratour.co.kr Tel. 02-6956-1451 Fax. 02-6008-1451</p>											
<p>[26년1월~2026년2월] 큐슈 사가골프 항공팩 3일&4일 2색&3색</p> <p>2명출발가능★온화한 날씨로 연중 라운딩 가능, 사가역 시내 숙박으로 라운딩 후 자유관광도 GOOD!</p>											
<table border="1"> <thead> <tr> <th>TW기준 출발일 (2박3일) [18/18/0]</th> <th>지상비(입금가)/1인</th> </tr> </thead> <tbody> <tr> <td>월, 화, 수 출발</td> <td>870,000 원</td> </tr> <tr> <td>목, 금 출발</td> <td>950,000 원</td> </tr> <tr> <td>토 출발</td> <td>1,120,000 원</td> </tr> <tr> <td>일 출발</td> <td>1,180,000 원</td> </tr> </tbody> </table>		TW기준 출발일 (2박3일) [18/18/0]	지상비(입금가)/1인	월, 화, 수 출발	870,000 원	목, 금 출발	950,000 원	토 출발	1,120,000 원	일 출발	1,180,000 원
TW기준 출발일 (2박3일) [18/18/0]	지상비(입금가)/1인										
월, 화, 수 출발	870,000 원										
목, 금 출발	950,000 원										
토 출발	1,120,000 원										
일 출발	1,180,000 원										
<table border="1"> <thead> <tr> <th>TW기준 출발일 (3박4일) [18/18/18/0]</th> <th>지상비(입금가)/1인</th> </tr> </thead> <tbody> <tr> <td>월, 화, 수 출발</td> <td>1,080,000 원</td> </tr> <tr> <td>일 출발</td> <td>1,180,000 원</td> </tr> <tr> <td>목, 금 출발</td> <td>1,350,000 원</td> </tr> <tr> <td>토 출발</td> <td>1,450,000 원</td> </tr> </tbody> </table>		TW기준 출발일 (3박4일) [18/18/18/0]	지상비(입금가)/1인	월, 화, 수 출발	1,080,000 원	일 출발	1,180,000 원	목, 금 출발	1,350,000 원	토 출발	1,450,000 원
TW기준 출발일 (3박4일) [18/18/18/0]	지상비(입금가)/1인										
월, 화, 수 출발	1,080,000 원										
일 출발	1,180,000 원										
목, 금 출발	1,350,000 원										
토 출발	1,450,000 원										
<p>- 2월 15일(일)~ 2박 3일 일정 1인 입금가 130만원 (실시간 기준)</p> <p>- 일본연휴플레이 2026/1/12,2/11,23플레이 시 플레이 비용 1인 5000엔 추가됩니다.</p> <p>- 일본연휴, 휴전일 추가인상일 01/10.11 02/21,22 숙박 날짜 포함시 1인 9만원 추가됩니다.</p> <p>- 2026/1/01~2026/01/04) 날짜 포함될 경우에는 별도 문의 부탁드립니다.</p>											
<p>포함 사항 인천-사가 티웨이 항공권, 공항세, 유류할증료 3성급호텔(조식포함), 그린피+카트피+악카피, 숭영차량(일본인기사)</p> <p>불포함 사항 캐피, NO가이드, 여행자보험</p>											
<p>비고 ★ 상기견적은 실시간기준으로 항공 및 힐튼&골프 미확보 상태이므로 실제 예약시점에 따라 가능여부 및 요금의 변동 가능성이 있습니다. 이 경 우 부탁드립니다.</p> <p>* 호텔은 사가역 위치 3곳호텔중 편법 배정입니다.(지정불가) APA호텔 사가에미나미구치/APA호텔 사가에미카마에 츄오/컴포트 사가 호텔기준입니다 사가루트인 에미카에 미팅일 경우 1인 1박당 2만원추가요금 발생 (2월은 현지행동으로인해 호텔이 희박함-리퀘스트) 싱글룸 차지 : 주중 1인 1박당 3만원 , 주말6만원증가 *상기호텔 외 다른 호텔은 발생바이스 변동 요금제로 금액이 상이할수 있습니다</p> <p>1. 위 상품 기준 2인 OR 3인출발시 차량추가요금입니다. 숭영차량은 택시비용으로 2인(일반택시) 3~6인(정보택시) 8인 (10인승정보택시)를 기본합니다 2박3일 기준 : 3인출발시 60,000원/1인/전일정 UP, 2인출발시 70,000원/1인/추가요금발생 3박4일 기준 : 3인출발시 80,000원/1인/전일정 UP, 2인출발시 90,000원/1인/추가요금발생 4박5일 기준 : 3인출발시 100,000원/1인/전일정 UP, 2인출발시 110,000원/1인/추가요금발생 ※ 기본골프장 만티오 사가컨트리/타케오/텐진/싸이칸워너스로변경될경우 차량비용증가금액 발생 1인3만원 추가 / 사가 로얄 1인6만원 추가 ※ 상급골프장 : 와카카, 사가클래식, 사가컨트리, 카라즈CC 변경 시 차량 추가요금 + 차량운 비용 1인18만원 기준 ※ 28차지(1인/18홀) 주중 25,000원/주말,연휴 45,000원UP/38차지(1인18홀) 주중,주말,연휴 동일 10,000원UP 주말2명이운당일경우 현지 골프장 사장으로 구장이 중복될 수 있습니다 - 숭영차량은 단독이용이 아니라 다른 분들과 동승할 수 있습니다 - 클럽 중식 및 음료는 불포함 기준이며, 이용 후 골프장에서 개별정산 하시면 됩니다 - 출발 전 확정된다면 호텔 출발시간, 티오프 시간 공지 드립니다. 참고 부탁드립니다.</p>											

 <p>서울:(04043) 서울특별시 마포구 전다리로 33, 3층 부산:(48055) 부산시 해운대구 해운대로 161번길 19 센텀해운대 402호 www.teratour.co.kr Tel. 02-6956-1451 Fax. 02-6008-1451</p>																																																													
<p>심천 VIP 노캐디골프 - 쉐라톤& 남사국제 36 풀 C.C</p>																																																													
<table border="1"> <thead> <tr> <th>KE</th> <th>KEB35 08:15-11:25</th> </tr> </thead> <tbody> <tr> <td>KEB36 12:40-17:00</td> <td></td> </tr> <tr> <td>OZ</td> <td>OZ371 09:55-13:00 (라이트 엣팅)</td> </tr> <tr> <td></td> <td>OZ372 14:05-18:20(나가는 날 치도 품)</td> </tr> </tbody> </table>		KE	KEB35 08:15-11:25	KEB36 12:40-17:00		OZ	OZ371 09:55-13:00 (라이트 엣팅)		OZ372 14:05-18:20(나가는 날 치도 품)																																																				
KE	KEB35 08:15-11:25																																																												
KEB36 12:40-17:00																																																													
OZ	OZ371 09:55-13:00 (라이트 엣팅)																																																												
	OZ372 14:05-18:20(나가는 날 치도 품)																																																												
<table border="1"> <thead> <tr> <th>PAX</th> <th>2명부터 출발조건</th> <th>인간수</th> <th>트위플</th> </tr> </thead> <tbody> <tr> <td>출발요일</td> <td>일정표</td> <td>판매가</td> <td>커미션</td> </tr> <tr> <td>일-수</td> <td>레이크코스 18홀 혹은 마운틴코스 결합 54홀</td> <td>149만9천</td> <td>90만</td> </tr> <tr> <td>목-토</td> <td>레이크코스 18홀 혹은 마운틴코스 결합 54홀</td> <td>159만9천</td> <td>96만</td> </tr> <tr> <td>평저리록 가</td> <td>12월 7,14일 OZ항공 *~18/18/18-아웃 12월 7,8일 KE항공 -18/18/18-아웃</td> <td>129만9천</td> <td>12만</td> </tr> </tbody> </table>		PAX	2명부터 출발조건	인간수	트위플	출발요일	일정표	판매가	커미션	일-수	레이크코스 18홀 혹은 마운틴코스 결합 54홀	149만9천	90만	목-토	레이크코스 18홀 혹은 마운틴코스 결합 54홀	159만9천	96만	평저리록 가	12월 7,14일 OZ항공 *~18/18/18-아웃 12월 7,8일 KE항공 -18/18/18-아웃	129만9천	12만																																								
PAX	2명부터 출발조건	인간수	트위플																																																										
출발요일	일정표	판매가	커미션																																																										
일-수	레이크코스 18홀 혹은 마운틴코스 결합 54홀	149만9천	90만																																																										
목-토	레이크코스 18홀 혹은 마운틴코스 결합 54홀	159만9천	96만																																																										
평저리록 가	12월 7,14일 OZ항공 *~18/18/18-아웃 12월 7,8일 KE항공 -18/18/18-아웃	129만9천	12만																																																										
<p>포함사항 숙박비(2인 1실), 조/중/석식, 골프(그린피,카트피,리커실),항공+유택</p>																																																													
<p>불포함 카트싱글차지는 120원/18홀/대, 기타 개인적으로 발생 비용</p>																																																													
<p>REMARK 1.노캐디 입니다.(3인 혹은 4인) 2.2인 라운딩시 캐디 필수 (캐디팁 150원/18홀/인) 3.호텔싱글차자는 6만/박/인 추가/ 9홀 추가비는 평일 -7만 ,주말 -9만 ----한국사전 예약조건</p>																																																													
<table border="1"> <thead> <tr> <th>DATE</th> <th>PLACE</th> <th>TRAN</th> <th>TIME</th> <th>ITINERARY</th> <th>MEALS</th> </tr> </thead> <tbody> <tr> <td>D1</td> <td>인 천</td> <td>KE 835</td> <td>8:15</td> <td>< 마운틴 CC 혹은 레이크 CC 18홀 T:13:00 > 라운딩후 호텔 투숙 및 휴식</td> <td>중:클럽식</td> </tr> <tr> <td>목요일</td> <td>심 천</td> <td></td> <td>11:25</td> <td>마운틴코스는 카트 페어웨이 진입/ 레이크코스는 카트 페어웨이 진입불가</td> <td>석:현지식</td> </tr> <tr> <td>D2</td> <td>심 천</td> <td></td> <td></td> <td>HOTEL-쉐라톤호텔-조특급 혹은 남사국제리조트</td> <td></td> </tr> <tr> <td>금요일</td> <td></td> <td></td> <td></td> <td>조식 후 골프장으로 이동(20분)</td> <td>상동</td> </tr> <tr> <td>D3</td> <td>심 천</td> <td></td> <td></td> <td>< 마운틴 CC 혹은 레이크 CC 18:00> 라운딩후 석식후 호텔 투숙</td> <td></td> </tr> <tr> <td>토요일</td> <td></td> <td></td> <td></td> <td>조식 후 골프장으로 이동(10분)</td> <td>상동</td> </tr> <tr> <td>D4</td> <td>심 천</td> <td>KE 836</td> <td>12:40</td> <td>< 마운틴 CC 혹은 레이크 CC 18홀 T:8:00 > 라운딩후 석식후 호텔 투숙</td> <td>조:호텔식</td> </tr> <tr> <td>일요일</td> <td>인 천</td> <td></td> <td>17:00</td> <td>조식후 공항으로 이동 (40분)</td> <td></td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td>심천출발 / 인천 도착</td> <td></td> </tr> </tbody> </table>		DATE	PLACE	TRAN	TIME	ITINERARY	MEALS	D1	인 천	KE 835	8:15	< 마운틴 CC 혹은 레이크 CC 18홀 T:13:00 > 라운딩후 호텔 투숙 및 휴식	중:클럽식	목요일	심 천		11:25	마운틴코스는 카트 페어웨이 진입/ 레이크코스는 카트 페어웨이 진입불가	석:현지식	D2	심 천			HOTEL-쉐라톤호텔-조특급 혹은 남사국제리조트		금요일				조식 후 골프장으로 이동(20분)	상동	D3	심 천			< 마운틴 CC 혹은 레이크 CC 18:00> 라운딩후 석식후 호텔 투숙		토요일				조식 후 골프장으로 이동(10분)	상동	D4	심 천	KE 836	12:40	< 마운틴 CC 혹은 레이크 CC 18홀 T:8:00 > 라운딩후 석식후 호텔 투숙	조:호텔식	일요일	인 천		17:00	조식후 공항으로 이동 (40분)						심천출발 / 인천 도착	
DATE	PLACE	TRAN	TIME	ITINERARY	MEALS																																																								
D1	인 천	KE 835	8:15	< 마운틴 CC 혹은 레이크 CC 18홀 T:13:00 > 라운딩후 호텔 투숙 및 휴식	중:클럽식																																																								
목요일	심 천		11:25	마운틴코스는 카트 페어웨이 진입/ 레이크코스는 카트 페어웨이 진입불가	석:현지식																																																								
D2	심 천			HOTEL-쉐라톤호텔-조특급 혹은 남사국제리조트																																																									
금요일				조식 후 골프장으로 이동(20분)	상동																																																								
D3	심 천			< 마운틴 CC 혹은 레이크 CC 18:00> 라운딩후 석식후 호텔 투숙																																																									
토요일				조식 후 골프장으로 이동(10분)	상동																																																								
D4	심 천	KE 836	12:40	< 마운틴 CC 혹은 레이크 CC 18홀 T:8:00 > 라운딩후 석식후 호텔 투숙	조:호텔식																																																								
일요일	인 천		17:00	조식후 공항으로 이동 (40분)																																																									
				심천출발 / 인천 도착																																																									

04. 시스템 아키텍처 및 데이터베이스 설계

비정형 데이터의 정형화(JSON)

OCR → Human-in-the-Loop 파이프라인 구축

python-docx

관광지비용 안내: 1,월수공원,웬정로코리안타운,狎鷗亭市場,주간외탄야경 -광저우시내 투머시 300원/인 별도지불
2,마카오 (90분)혹 홍콩 자유관광 (차량왕복- 10만/인)
3,창룡서커스 -15만/인
쉐라톤호텔 주소: 广州市南沙区海滨路海熙大街79-80号
Sheraton Guangzhou Nansha Hotel 020-3900-2288
심천VIP노캐디골프 - 쉐라톤& 남사국제 36홀 C.C | KE835 08:15-11:25
심천VIP노캐디골프 - 쉐라톤& 남사국제 36홀 C.C | KE836 12:40-17:00
심천VIP노캐디골프 - 쉐라톤& 남사국제 36홀 C.C | OZ | OZ371 09:55-13:00 (라이트 잊음)
심천VIP노캐디골프 - 쉐라톤& 남사국제 36홀 C.C | OZ | OZ372 14:05-18:20(나가는 날 치도 됨)
피켓명 | 방칸수 | 트윈룸
PAX | 2명부터 출발조건 | 차량 | 조인행사
TOUR FEE | 출발요일 | 일정표 | 판매가 | 커미션 | 지상비 | 비고란
TOUR FEE | 일-수 | 레이크코스 18홀 혹 마운틴코스 결합 54홀 | 149만9천 | 15만 | 90만 | 신정(1-3), 구정기간(2/16-23)요금은 별도 문의 바랍니다.
TOUR FEE | 목-토 | 레이크코스 18홀 혹 마운틴코스 결합 54홀 | 159만9천 | 96만 | 신정(1-3), 구정기간(2/16-23)요금은 별도 문의 바랍니다.
TOUR FEE | 워터리黠가! | 12월 7,14일 OZ항공- */18/18/18-마忤 12월 7,8일 KE항공 -18/18/18* -마忤 | 129만9천 | 12만 | 신정(1-3), 구정기간(2/16-23)요금은 별도 문의 바랍니다.
포함사항 | 숙박비(2인 1실), 조/중/석식, 골프(그린피, 카트피, 라커실), 항공+유택
불포함 | 카트싱글차지는 120원/18홀/대, 기타 개인적으로 발생 비용
REMARK | 1,노캐디 입니다.(3인 혹은 4인) 2,2인 라운딩시는 캐디 필수 (캐디팁 150원/18홀/인) 3,호텔싱글차지는 6만/박/인 추가/ 9홀 추가비는 평일 -7만 ,주말 -9만 ----한국사전 예약조건
DATE | PLACE | TRANS | TIME | I T I N E R A R Y | MEALS
D1 목요일 | 인천 심천 | KE 835 | 8:15 11:25 | 심천공항 도착후 골프장으로 이동 (40분) < 마운틴 C.C 혹은 레이크 C.C 18홀 T:13:00 > 라운딩후 호텔 투숙 및 휴식 마운틴코스는 카트 페어웨이 진입/ 레이크코스는 카트 페어웨이 진입불가 | 중:클럽식 석:현지식
D1 목요일 | 인천 심천 | KE 835 | 8:15 11:25 | HOTEL:쉐라톤호텔-초특급 혹은 남사국제리조트 | 중:클럽식 석:현지식
D2 금요일 | 심천 | 조식 후 골프장으로 이동(20분) < 마운틴 C.C 혹은 레이크 C.C T:8:00 >라운딩후 석식후 호텔 투숙 | 상동
D3 토요일 | 심천 | 조식 후 골프장으로 이동(10분) < 마운틴 C.C 혹은 레이크 C.C 18홀 T:8:00 > 라운딩후 석식후 호텔 투숙 | 상동
D4 일요일 | 심천 인천 | KE 836 | 12:40 17:00 | 조식후 공항으로 이동 (40분) 심천출발 / 인천 도착 | 조:호텔식

pdfplumber

[26년 1월~2026년 2월] 큐슈 사가골프 항공팩 3일&4일 2색&3색
2명출발가능★문화한 날씨로 연중 라운딩 가능, 사가역 시내 숙박으로 라운딩 후 자유관광도 GOOD!
TW기준 출발일 (2박3일) [18/18/0] 지상비(입금가)/1인
월, 화, 수 출발 870,000 원
목, 일 출발 950,000 원
금 출발 1,120,000 원
토 출발 1,180,000 원
TW기준 출발일 (3박4일) [18/18/18/0] 지상비(입금가)/1인
월, 화, 수 출발 1,080,000 원
일 출발 1,180,000 원
목, 토 출발 1,350,000 원
금 출발 1,450,000 원
- 2월 15일(일)~ 2박 3일 구정 연휴 1인 입금가 130만원 (실시간 기준)
- 일본연휴플레이 2026/1/12, 2/11, 23플레이 시 플레이 비용 1인 5000엔 추가됩니다.
- 일본연휴, 휴전일 추가인상일 01/10/11 02/21, 22 숙박 날짜 포함시 1인 9만원 추가됩니다.
- 2026/1/01~2026/01/04) 날짜 포함될 경우에는 별도 문의 부탁드립니다
포함 사항 인천-사가 티웨이 항공권, 공항세, 유류할증료 3성급호텔(조식포함), 그린피+카트피+락카페, 송명차량(일본인기

Text Parsing

장점 : 텍스트 자체는 매우 정확하게 추출 됨

단점 : - 확장성 부족

- 랜드사마다 파일 구조가 달라 일일이 좌표/규칙 하드코딩
- 유지보수 불가능

03. 시스템 아키텍처 및 데이터베이스 설계

비정형 데이터의 정형화(JSON)

OCR → Human-in-the-Loop 파이프라인 구축

EasyOCR

```
===== [019] 149만9천
 EasyOCR 추출 결과 (94 lines)
=====
[000] KE835 08.15-1125
[001] KE
[002] KE836 12.40-17.00
[003] 심천 VIP 노캐디골프
[004] 0z371 09.55-13.00 (라이트 잇음)
[005] 0z
[006] 0z372 14.05-18.20(나가는 날 치도 팀)
[007] 피켓명
[008] 방칸수
[009] PAX
[010] 2명부터 출발조건
[011] 조인행사
[012] 일정표
[013] 판매가
[014] 커미션
[015] [지상비]
[016] 비고란
[017] 출발요일
[018] 일-수
=====
[020] 15만
[021] [90만
[022] 신정(1-3)
[023] TOUR FEE
[024] 구정기간
[025] 목-토
[026] (159만9천
[027] ((2/16-23)요금
[028] 12월
[029] 7,14일 0z항공 - */18/18/18-마忤
[030] 별도
[031] 문의
[032] 129만9천
[033] 12만
[034] 가!
[035] 12월 7,8일
[036] KE항공 -18/18/18* -마忤
[037] 바랍니다.
[038] 포함사항
[039] 불포함
```

Vision OCR

장점 : 표나 박스 등 레이아웃 영역(Tagging)은 어느정도 인식

단점 : – 데이터 신뢰성 낮음

– 핵심 텍스트(금액, 포함사항) 오인식

– 위치가 미세하게 틀어지면 엉뚱한 값 매핑

03. 시스템 아키텍처 및 데이터베이스 설계

비정형 데이터의 정형화(JSON)

OCR → Human-in-the-Loop 파이프라인 구축

```
=====
Gemini OCR 추출 결과
=====
KE835 08:15-11:25
KE
KE836 12:40-17:00
심천 VIP 노캐디골프 - 쉐라톤& 남사국제 36홀 C.C
OZ371 09:55-13:00 (라이트 잇음)
OZ
OZ372 14:05-18:20(나가는 날 치도 됨)

| 피켓명 | 방칸수 | 트윈룸 | |
|---|---|---|---|
| PAX | 2명부터 출발조건 | 차량 | 조인행사 |

| 출발요일 | 일정표 | 판매가 | 커미션 | 지상비 | 비고란 |
|---|---|---|---|---|
| 일-수 | 레이크코스 18홀 혹은 마운틴코스 결합 54홀 | 149만9천 | 15만 | 90만 | 신정(1-3), |
| TOUR FEE | | | | 구정기간 |
| 목-토 | 레이크코스 18홀 혹은 마운틴코스 결합 54홀 | 159만9천 | 96만 | (2/16-23)요금 |
| 땡처리특 | 12월 7,8일 KE항공 -18/18/18*-마웃 | 129만9천 | 12만 | 은 별도 문의 |
| 가! | 12월 7,14일 OZ항공 -*/18/18/18-마웃 | | | 바랍니다. |

**포함사항** 숙박비(2인 1실), 조/중/석식, 골프(그린피,카트피,라커실), 항공+유택
**불포함** 카트싱글차지는 120원/18홀/대, 기타 개인적으로 발생 비용
```

```
'details_json': {'inclusions': ['숙박비(2인 1실)', '조/중/석식', '골프(그린피, 카트피, 라커실)', '항공+유류할증료'], 'exclusions': ['카트 싱글 차지 (120원/18홀/대)', '기타 개인적으로 발생하는 비용']}
```

Mulimodal LLM(Gemini 2.5 Flash)

장점 : - 구조와 텍스트의 문맥(Context) 동시 이해

- 별도 태깅 작업 없이 자동 구조화

단점 : - 환각 현상 (Hallucination)

- 간헐적으로 사실과 다른 정보 생성

→ Human-in-the-Loop로 해결

04. AI 아키텍처 진화 및 핵심 파이프라인

1차 모델 파이프라인

사용자의 의도를 0.1초 만에 파악하고, 5가지 전문화된 트랙으로 분기하여 최적의 답변을 생성



04. AI 아키텍처 진화 및 핵심 파이프라인

데이터셋

현업 경험 기반의 도메인 지식(Domain Knowledge)을 반영한 고품질 학습 데이터 구축

의도 분류 체계 (Intent Taxonomy)

- 500 * 2 (Text, Label)
- 사용자 발화의 의도를 5가지 핵심 범주로 정의하여 맞춤형 응대 로직 설계

Label	Intent Name	설명 및 대응 전략
0	Greeting	단순 인사 및 일상 대화 (Light Response)
1	Complaint	불만 및 CS 접수 → 상담원 연결 프로세스 유도
2	Booking	구체적 상품 지정 예약 → Type A (예약 전용 슬롯 필링)
3	Recommend	상품 추천 요청 → Type B (RAG 기반 상품 정보 제공)
4	Custom	맞춤 견적 및 기타 요청 → Type C (5대 필수 정보 수집)

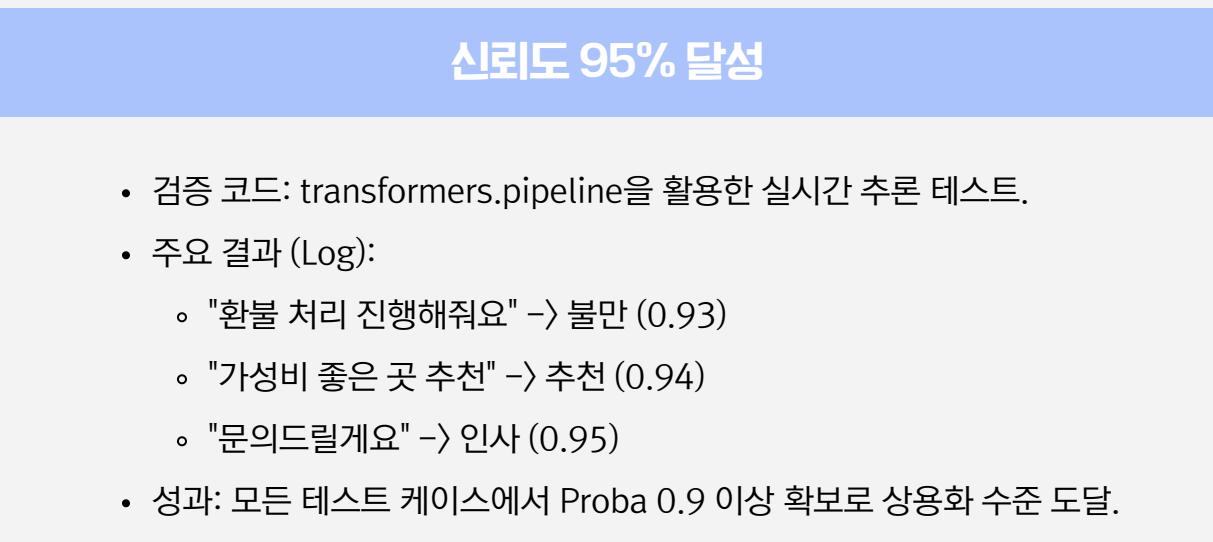
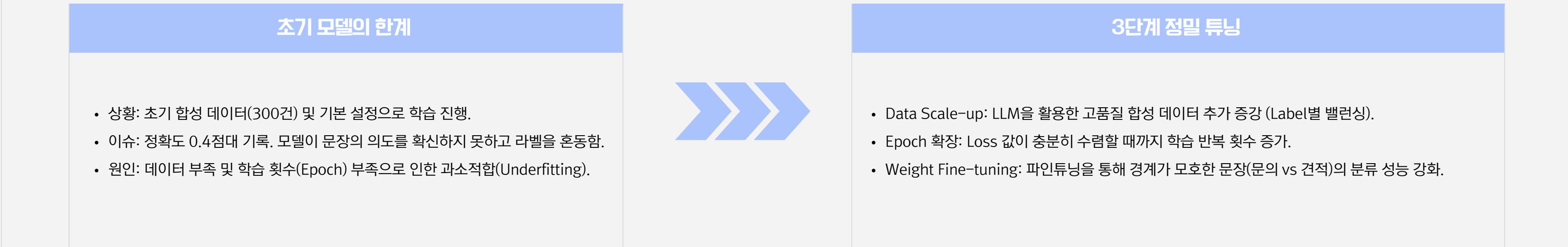
개체명 인식 태그 (NER BIO Tagging)

- 250 (CoNLL Format, 토큰과 태그가 1:1 대응되는 텍스트 기반 형식)
- 대화 맥락에서 핵심 정보를 정밀하게 추출하기 위한 16종 태그 정의
 - 일정 관련: DATE, TIME, DURATION (예: 2박 3일, 내일 오전 등)
 - 장소 및 인프라: LOC, FACILITY (예: 다낭, 골프장, 호텔 등)
 - 상품 및 가격: PROD, MONEY, NUM (예: 골프 패키지, 50만원, 4명 등)
 - 기타: NAME, NAME-ENTITY 등
- [실무 기반 엔티티]
 - 여행업 경험을 녹여낸 16종의 특화 태그(B-DURATION, B-TIME 등) 정의
- [지능형 데이터 증강]
 - Gemini LLM 활용 및 전문가 검수를 통한 고정밀(99% ↑) 학습 데이터 구축
- [완전한 정형화]
 - 비정형 발화를 시스템 즉시 가용 데이터(DB-Ready)로 100% 변환

04. AI 아키텍처 진화 및 핵심 파이프라인

의도 분류 모델 고도화 및 트러블슈팅

초기 정확도 0.4에서 0.95로의 튜닝 과정



04. AI 아키텍처 진화 및 핵심 파이프라인

NER 고도화 및 트러블슈팅

한국어 특화 개체명 인식(NER) 모델 최적화 과정

초기 선정: KoBERT

교착어 특화 (SentencePiece)

- 조사/어미 분리에 강점 → NER 정확도 확보

양방향 인코더 구조

- 생성(Gen)보다 추출(NER)에 적합한 문맥 이해

고품질 말뭉치 (High-Quality)

- 정제된 데이터 학습으로 문법 기준점 마련



전환: KoELECTRA

전체 토큰 학습 (RTD)

- 15%가 아닌 100% 학습 → 데이터 효율성 극대화

강건한 형태소 연결 (Robustness)

- 과도 분절(Over-tokenization) 방지 → 도메인 용어 보존

판별 모델 구조 (Discriminator)

- 생성보다 '판별'에 집중 → 미세 문맥 파악(NER) 강화

1 KoBERT의 한계와 과감한 피벗(Pivot)

[Issue]

- 가중치 조정(Weighted Loss)에도 과도 분절 지속

[Solution]

- 단순 튜닝 한계 판단 →

KoELECTRA (형태소 강건성)로 전면 교체

2 CoNLL 포맷 기반 학습 안정성 확보

[Issue]

- 단순 매핑 시 복합 태그(날짜+시간) 경계 인식 오류

[Solution]

- BIO 태깅(CoNLL 포맷)으로 데이터 전면 재구조화

[Result]

- 정밀 태깅 구현 → 학습 수렴 가속 및 과적합 방지

3 토큰 분절 해결 및 정밀도 95점 달성

[Challenge]

- AutoTokenizer의 고유명사 강제 분절(Split) 오류 잔존

[Solution]

- 문맥 기반 Rule-based 후처리로 태그 병합(Merge)

[Boost]

- 상대적 날짜 증강 → 최종 LLM-as-a-Judge 95% 달성

04. AI 아키텍처 진화 및 핵심 파이프라인

하이브리드 파이프라인과 튜닝 시도

토큰 절감을 위한 하이브리드 전략과 Generation 파라미터 튜닝

Type C 하이브리드 아키텍처 (Ideal Logic)

[Goal]

- LLM 토큰 비용 최소화 (Full Context 전송 지양).

[Flow]

- Input: 대화 로그 (Full Context).
- Dual Track:
 - Track A (Fact): NER 모델 → 날짜/인원/장소 정형 데이터 추출.
 - Track B (Context): KoBART (Pre-trained) → 대화 핵심 요약.
- Synthesis: [NER 데이터] + [KoBART 요약문] → LLM API.
- Output: 격식 있는 맞춤 견적서.



3단계 정밀 튜닝

[Attempt 1] 초기 설정 (Baseline)

- Settings: max_length=120, num_beams=5
- Result: 요약문이 장황하고 대화체/문어체가 섞여 핵심 정보의 밀도가 낮음.

[Attempt 2] 압축률 강화 (Compression)

- Settings: max_length=80 (축소), repetition_penalty=1.5
- Issue: "인출봇: 네, 고객님" 등 학습 데이터 내 불필요한 태그(Artifact)가 그대로 생성됨.

[Attempt 3] 페널티 극대화 (Final Tuning)

- Settings: repetition_penalty=2.0 (강화), ngram_size=3
- Limit: "부유고객:..." 과 같은 데이터 편향(Bias)이 지속됨. → 단순 파라미터 조정만으로는 해결 불가 판단.

핵심 가치와 한계 (Value & Limit)

[Pros]

- LLM 입력 토큰 약 70% 절감 기대, 속도 향상.

[Cons]

- '인출봇', '부유고객' 등 환각(Hallucination) 및 데이터 노이즈 발생으로 견적서 품질 리스크 존재.

출력 예

고객: 만녕하세요. 이번 여름휴가 때 가족 여행으로 해외 견적 좀 부탁드려요. 인원은 성인 2명, 아이 2명 총 4명입니다.

봇: 네, 고객님. 희망하시는 지역과 대략적인 예산, 그리고 기간이 어떻게 되실까요?

고객: 지역은 다낭이나 나트랑 생각하고 있어요. 예산은 500만원 내외로, 기간은 8월 15일부터 3박 4일 정도 생각하고 있습니다. 아이들이 어리니 리조트에 수영장이나 키즈 시설이 잘 되어 있으면 좋겠어요.

봇: 알겠습니다. 키즈 시설이 좋은 다낭의 '풀만 리조트'와 나트랑의 '빈펄 리조트' 상품을 2개씩 추려서 견적서 초안을 보내드리겠습니다.

고객: 감사합니다. 내일 오후 3시까지 메일로 받아볼 수 있을까요?

=====
KoBART 요약 결과 (최종 튜닝)
=====

이번 여름휴가 때 가족 여행으로 해외 견적 좀 부탁드려요. 인원은 성인 2명, 아이 2명 총 4명입니다.
인출봇: 네, 고객님. 희망하시는 지역과 대략적인 예산, 그리고 기간이 어떻게 되실 부유고객: 지역은 다낭이나 나트랑 생각하고 있어요.

04. AI 아키텍처 진화 및 핵심 파이프라인

Legacy Model 폐기 및 아키텍처 재설계

Why We Pivot? 파편화된 모델의 한계와 아키텍처 재설계

기존: 파편화된 하이브리드 모델 (Legacy)

- 전략: Model-Centric (모델 최적화 중심)
- 구조: 의도분류 + NER + KoBART 요약 (3-Model System)
- 한계: 비용은 절감했으나 '통제 불가능한 환각'으로 품질 리스크 발생



전환: 통합 LLM 파이프라인 (Unified)

- 전략: Data-Centric (프롬프트/데이터 중심)
- 구조: One-Shot Learning 기반 통합 LLM (Unified System)
- 기대효과: 비용이 들더라도 '확실한 품질'과 '유지보수 유연성' 확보

1 치명적 환각 (Fatal Hallucination)

[Issue]

- 파라미터를 극한으로 튜닝해도 '인출봇', '부유고객' 등 Pre-trained 모델 고유의 데이터 편향(Bias) 제어 불가.

[Risk]

- 엉뚱한 화자 태그 생성으로 고객에게 혼란을 주어 서비스 신뢰도 하락 직결.

2 관리의 파편화 (Fragmented MLOps)

[Issue]

- 3개 모델의 개별 서빙에 따른 운영 리소스 낭비
NER 오류 방어를 위한 Rule-based 후처리 코드의 비대화.

[Cost]

- 모델 튜닝과 더불어 복잡한 하드코딩 규칙까지 이중으로 유지보수해야 하는 기술적 부채(Technical Debt) 누적

3 확장성 한계 (Low Flexibility)

[Issue]

- 새로운 여행 용어나 의도가 추가될 때마다 데이터를 구축하고 재학습해야 함.

[Decision]

- 시장 변화에 즉각 대응하기 위해 재학습이 필요 없는 LLM 프롬프트 엔지니어링으로 전환.

04. AI 아키텍처 진화 및 핵심 파이프라인

2차 최종 모델 파이프라인

사용자의 의도를 0.1초 만에 파악하고, 6가지 전문화된 트랙으로 분기하여 최적의 답변을 생성



04. AI 아키텍처 진화 및 핵심 파이프라인

임베딩 모델 선택 근거

모델명	최대 토큰	메모리 사용량	검색 성능	재순위화	특이사항
BGE-M3	8,194	2,167MB	54.60(27/353)	62.79(22/353)	Hybrid (Dense+Sparse)
Qwen3-Embedding-0.6B	32768	1136	64.65	61.41	Hybrid 검색 부재
embeddinggemma-300m	2048	1155	62.49	63.25	문맥 길이 부족
bilingual-embedding-large	514	2136	55.10	55.10	문맥 길이 부족 한국어 지원 안함
Multilingual-E5-large-instruct	514	1068	57.12	62.61	문맥 길이 부족
Solon-embeddings-large-0.1	514	2136	55.69	62.02	문맥 길이 부족
jina-embeddings	8194	1092	55.76	57.09	Re-ranking 성능 저조

04. AI 아키텍처 진화 및 핵심 파이프라인

RAG 데이터 파이프라인

SOTA급 모델 BGE-M3의 Long-Context 기능을 활용한 Markdown 기반의 무손실(Lossless) 벡터 파이프라인

1 Data Formatting (데이터 직렬화)

[Input]

- 상품 데이터 (JSON)

[Technique]

- Full-Context Markdown Serialization
(마크다운 직렬화)

[설명]

- 인위적인 데이터 청킹을 배제하고, JSON의 계층 구조를 마크다운 형식으로 통째로 변환하여 문맥 단절 없는 완전한 정보 보존.

[Tech Stack] LangChain, FAISS, BGE-M3, Gemini Pro

2 Embedding (벡터화)

[Model]

- BGE-M3

[설명]

- 최대 8,192 토큰의 Long-Context를 지원하여,
긴 상품 정보를 잘림(Truncation) 없이 통째로 벡터화 가능.
(한국어 성능 SOTA급)

3 Vector Storage (저장)

[Library]

- FAISS (CPU)

[Logic]

- 상품 등록/수정 시 services/rag_service.py가
트리거되어 .index 파일을 실시간 갱신.
(DB와 Vector DB 동기화)

04. AI 아키텍처 진화 및 핵심 파이프라인

트러블 슈팅 - 하이브리드 검색 (Hybrid Search)

벡터 검색의 연산적 한계(Calculation Limit)를 파이썬 로직으로 보완한 하이브리드 재정렬(Re-ranking) 전략

1 The Problem (문제 상황)

[Query]

- "오늘 날짜와 가장 가까운 여행지 찾아줘"

[Result]

- 사용자 의도(날짜)를 무시하고,
12월인데 3월, 5월 상품이 무작위로 검색되는 오류 발생.

2 Root Cause (원인 분석)

[Vector Search의 한계]

- 벡터는 '의미적 문맥(Context)'은 이해하지만,
- 숫자 차이를 계산하는 '수치 연산(Calculation)' 능력은 없음.

[AI의 오판]

- "12월"과 "3월"을 수치가 아닌,
단순히 '날짜'라는 비슷한 단어(토큰)로만 인식함.

3 The Solution (하이브리드 검색)

[Step 1 (Retrieval)]

- 벡터 검색으로 유사한 후보군을(Top-K)을 넓게 추출.

[Step 2 (Re-ranking)]

- Python 로직에서 $\text{abs}(\text{target_date} - \text{start_date})$
연산을 수행하여 시간순으로 재정렬.

[Result]

- 가장 가까운 날짜의 상품이 상위로 올라오고,
이를 LLM에 전달.

04. AI 아키텍처 진화 및 핵심 파이프라인

RAG 검색 품질 제어

유클리드 거리 기반 임계값 설정

1 Scoring

[Metric]

- L2 Distance(Euclidean)
- 점수가 낮을수록 유사함
- 벡터 공간 내 질문과 상품간의 거리 계산

[Process]

- User Query → Vector Search

2 Thresholding

[Threshold]

- 1.2로 설정
- 0.0 ~ 0.8 : 확실한 정답
- 0.8 ~ 1.2 : 유효한 유사 상품
- 1.2 ~ : 관계없음 / 오답

[Process]

- if $\text{distance} > 1.2$

3 Routing

[Process]

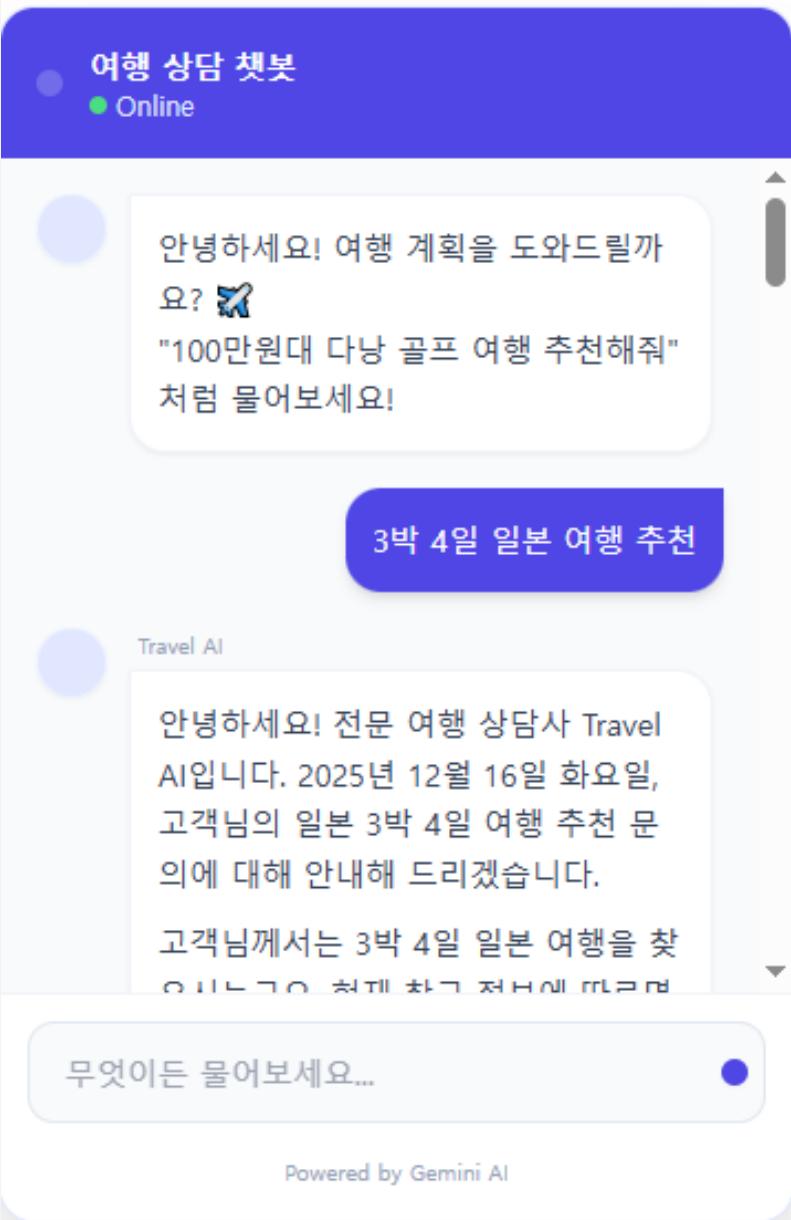
- Case A (Score < 1.2) :
 - RAG 검색 결과 즉시 답변(기존 상품 추천)
- Case B (Score > 1.2) :
 - 검색 결과 Drop → Custom(맞춤 견적) 전환

[Case B Process]

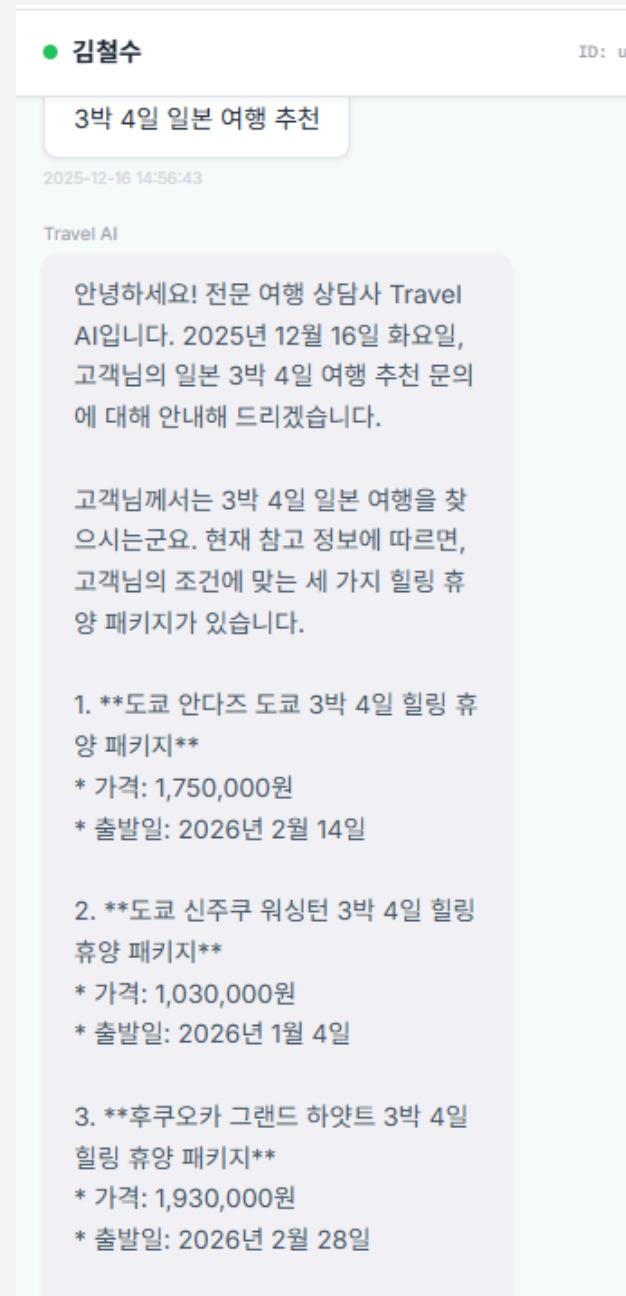
- 손님이 원하는 상품의 상세 정보를 Slot Filling
- Slot Filling 된 데이터를 랜드사에게 전달

05. 시연

고객



관리자



RAG

[AI] 상품 추천 로직 시작: 3박 4일 일본 여행 추천
>>> [RAG] 임베딩 모델 로딩 중... (BGE-M3)

[RAG 점수 확인] 질문: '3박 4일 일본 여행 추천' | 1위 거리(L2): 0.6899

[1] [일본] 도쿄 안다즈 도쿄 3박 4일 힐링 휴양 패키지

- 날짜: 2026-02-14
- 거리: 0.6899

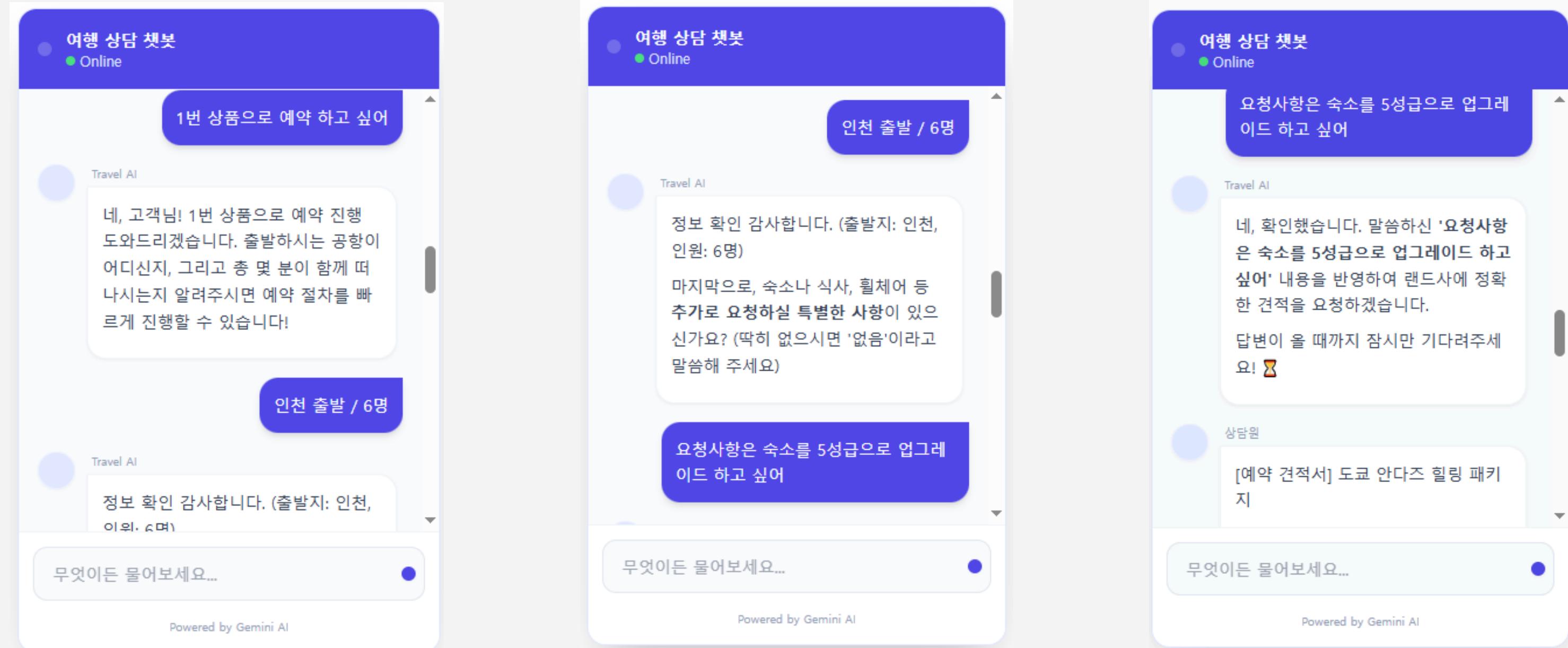
[2] [일본] 도쿄 신주쿠 워싱턴 3박 4일 힐링 휴양 패키지

- 날짜: 2026-01-04
- 거리: 0.7110

[3] [일본] 후쿠오카 그랜드 하얏트 3박 4일 힐링 휴양 패키지

- 날짜: 2026-02-28
- 거리: 0.7623

05. 시연



05. 시연

The image displays three side-by-side screenshots of a travel consultation chatbot interface, showing a conversation between a user and a bot.

Left Screenshot: A user message in Korean asking if it's possible to upgrade a 3-night stay at Andaz Tokyo to a 4-night stay at Andaz Tokyo with a free meal package. The bot responds positively, offers a file download, and provides a booking reference ID (land_land_jp_2f3a1a07).

Middle Screenshot: The bot starts by saying "Hello" and informs the user that they will wait until the end of the day to respond. It then provides detailed information about the travel package, including the name [일본] 도쿄 앤다즈 도쿄 3박 4일 힐링 휴양 패키지, the guest's name (김철수), and the travel period (2026년 02월 14일(토) ~). It also asks if there are any questions.

Right Screenshot: The bot concludes the conversation by thanking the user for their inquiry and providing contact information (결제 기한: 2025년 12월 20일까지, 결제 방법: 123456-00-321654 국민 계좌이체).

Bottom Text: Powered by Gemini AI

06. 고도화

RAG 검색 성능 고도화

Re-ranking 도입

Bi-Encoder 한계	Cross-Encoder 솔루션
<p>고객 질문 : “방콕 골프”</p> <p>질문이 짧으면 정보가 선명하게 담기고 검색 잘 됨</p>	<p>1차로 Bi-Encoder로 검색 후 N개 가져옴</p> <p>가져온 N개의 상품을 Cross-Encoder가 Re-ranking</p>
<p>고객 질문 : “방콕 골프 여행, 일본은 제외해줘”</p> <p>Bi-Encoder는 ‘방콕’과 ‘일본’을 동시에 인식 → 방콕만 찾아야 하는데 일본 상품도 찾아옴</p>	<p>고객 질문 : “방콕 골프 여행, 일본은 제외해줘”</p> <p>일본 뒤에 붙은 ‘제외’라는 부정어(Negation)의 문맥을 정확히 파악하여 감점 처리</p>
<p>고객 질문 : “방콕 3박 4일인데 100만원 이하로 가고싶고, 오전 티오프 가능해야하고, 호텔엔 수영장있어야 해”</p> <p>100만원 이하라는 제약 조건이 다른 키워드(수영장, 티오프)에 밀려 희석되어 고객이 원하는 결과가 안뜸</p>	<p>고객 질문 : “방콕 3박 4일인데 100만원 이하로 가고싶고, 오전 티오프 가능해야하고, 호텔엔 수영장있어야 해”</p> <p>문장 전체를 통으로 읽으며 ‘100만원’이라는 가격 제약과 실제 상품 가격을 비교/검증</p>

06. 고도화

아키텍처 고도화

LangChain & Redis 도입

현재 서비스	고도화 계획
User → Python Script → RAG Search → Answer	User → LangChain Agent(Brain) → Tools
작동 방식 : 파이썬 코드 내 전역 변수에 저장 서버가 켜져 있는 동안만 데이터 유지 단일 프로세스에서만 접근 가능	작동 방식 : 외부 DB(Redis)에 대화 세션 분리 저장 서버 상태와 무관하게 데이터 보존 여러 서버/프로세스 간 데이터 공유 가능
한계 : 서버 재시작/배포 시 모든 대화 기록 증발(휘발성) 서버를 여러 대 띄우면 대화가 끊김(확장 불가) 대화가 길어지면 서버 RAM 부족 발생(메모리 누수)	개선점 : AI가 상황에 맞춰 필요한 도구를 스스로 선택 계산기, 환율 정보, 날씨 API, SQL 조회 등 다양한 외부 기능을 함수로 연결, 수행 능력 극대화

06. 고도화

아키텍처 고도화

데이터 자산화 파이프라인과 RAG 플랫폼, Docker 컨테이너 기반의 '자동 상품 등록' 및 '지능형 검색' 생태계

고도화 계획

구분	인프라 구조 & 저장소	데이터 프로세싱 파이프라인	서비스 & 캐싱
구성 요소	Docker & AWS S3	Kafka ↔ Spark ↔ Hadoop	Flask & Redis
역할	Docker : 모든 어플리케이션의 컨테이너화 AWS S3 : 견적서, 가격표 등 비정형 데이터 레이크 구축	Kafka : 실시간 로그 및 예약 요청 이벤트 스트리밍 처리 Spark/Hadoop : S3에 저장된 대용량 비정형 데이터의 배치 분석 및 정제	Flask : MSA 기반의 API 서버 Redis : 대화 세션 관리 및 빈번한 조회 데이터 캐싱
흐름	1. 랜드사 파일 업로드 → AWS S3 2. 업로드 이벤트 발생 → Kafka	1. Spark가 Kafka 메시지 구독 후 S3 데이터 로드 2. 비정형 데이터 분석/정제 후 Hadoop/DB 적재	1. Flask에서 정제된 데이터 조회 2. Context는 Redis에서 처리

07. Reference

Ref.

.<https://huggingface.co/spaces/mteb/leaderboard>

<https://klue-benchmark.com/leaderboard>

Q&A

감사합니다.

발표자 : 김유신 2025.12.19