



PoliSight

(정치Politics + 통찰Insight + 시각Sight)
AI 기반 국회 정보 단순화 디지털 플랫폼

2팀 Outliers - 강원빈, 김예원, 김은채, 김호연, 박지아, 이준호

Contents

1. 문제정의
2. 솔루션 개요
3. 서비스 시연
4. 구성 및 역할
5. 구현
6. 확장성
7. 자체 평가
8. Appendix

1. 문제 정의: 국회의원의 활동, 얼마나 알고 계신가요?

- 대다수 국민은 국회의원의 활동 내용을 알기 어려움
- 회의록·표결 정보는 있지만 산발되어있고 정리된 내용이 없어 한 눈에 파악이 어려움
- 정보 비대칭은 정치 무관심과 불신으로 이어질 수 있음



2. 솔루션 개요: 국회 활동의 모든 것을 한 눈에

“국민이 쉽게 이해할 수 있도록
공공 API 기반으로 국회 데이터를 수집하고, 시각화하여 제공합니다.”

직관적인 정보 접근



정보 정제 및 시각화



데이터 수집

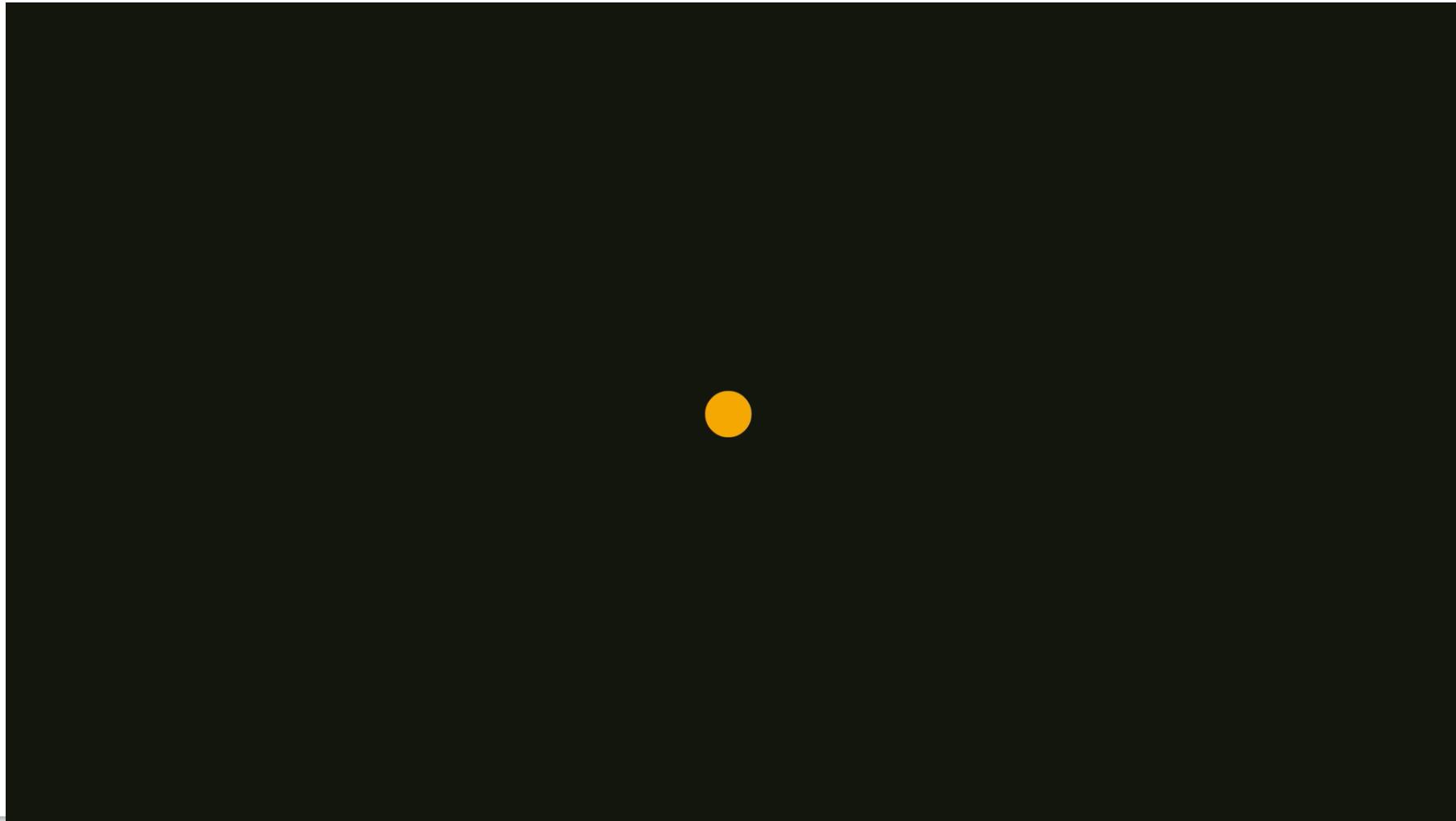


- 누구나 쉽게 의정 활동을 확인 가능
- 정보 소외 계층을 배려한 서비스 설계(TTS)

- 방대한 데이터를 요약
- 사용자 중심의 시각화
- 개인적 관심 기반 탐색

- 회의록, 표결 정보, 의원 정보 등 국회 데이터 수집
- 실시간 업데이트

3. 서비스 시연



4. 구성 및 역할



김호연

PM/BA

Data, 정책, 윤리



박지아

BA/ML Engineer

Machine Learning



김은채

PO/PL

Data, Serving, INFRA



강원빈

DevOps/AI Engineer

Model Training, Serving



이준호

AI Engineer/Devops

RAG, Data



김예원

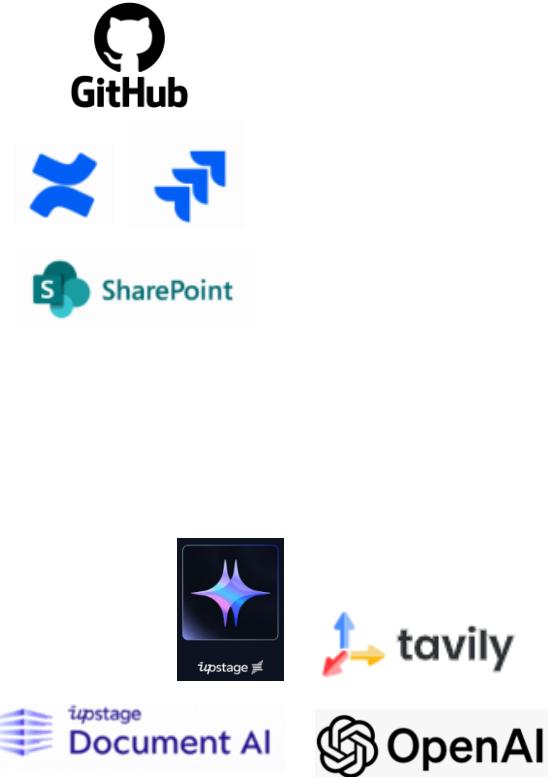
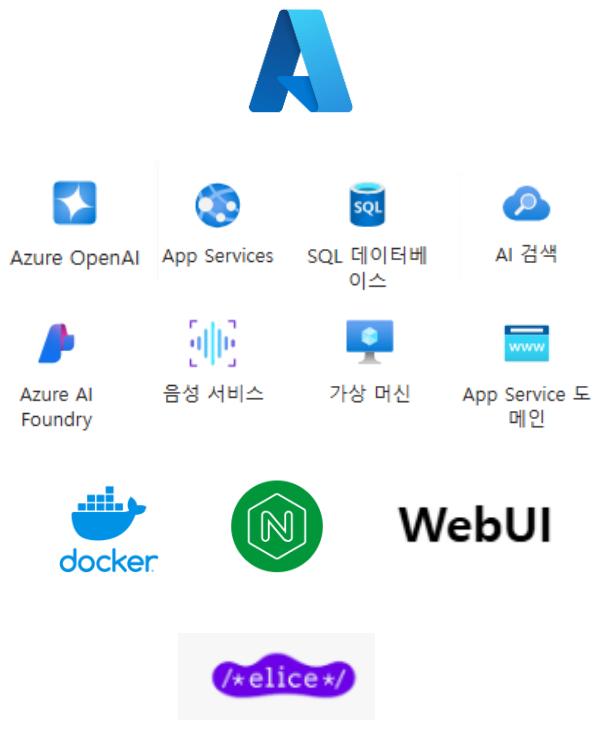
Backend/Data-Engineer

INFRA, Data

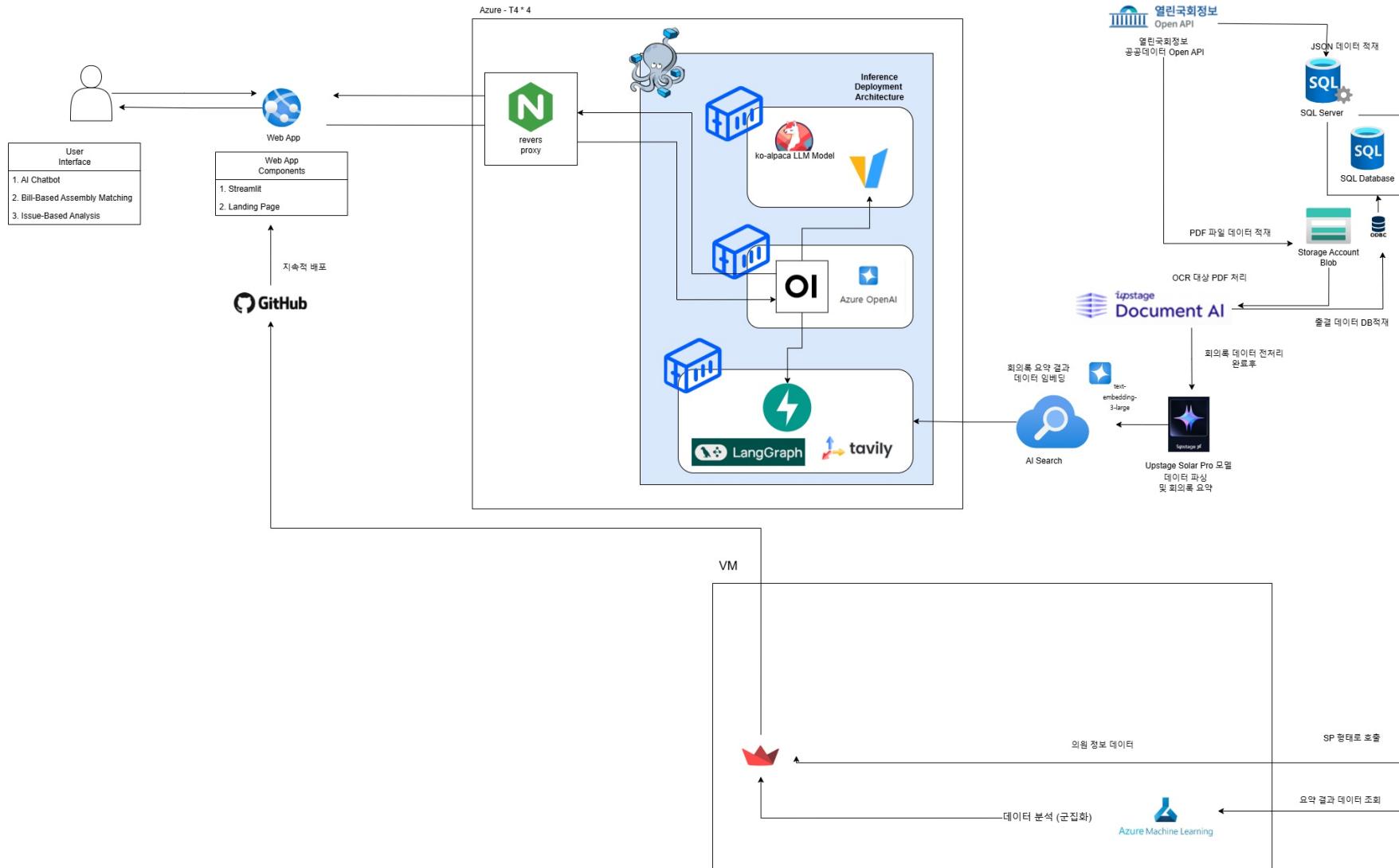
5. 구현: 작업 일정

	Phase0	Phase1				
일정	주제선정0	주제선정1	MVP1	MVP2	MVP3	MVP4
	06/19-22	6/23-24	6/25-26	6/27	6/28	6/29
	4Days	2Days	2Days	1Day	1Day	1Day
MVP		미정	1) 각 국회의원 별 발의안 군집화, 시각화 2) 긴 발의안을 축약해 의견 정리	국회의원 발의 내용 시각화 부분 음성 전달 구현	RAG 이용한 전체 페이지에 대한 챗봇 구축	챗봇 사용성 향상 (TTS)
작업	데이터셋 PoC Test	요구사항 정의서 데이터셋 구성 경쟁사 분석 인프라구성	데이터 처리 (DB적재 및 OCR 작업) GPU 세팅 ML 프로젝트 진행	LLM 모델 학습 Naïve RAG Azure Speech	Modular RAG WebUI & Streamlit 구성	인프라 보안 정비 Front 정리 PPT 발표자료 작성

5-1. Infra/Tool

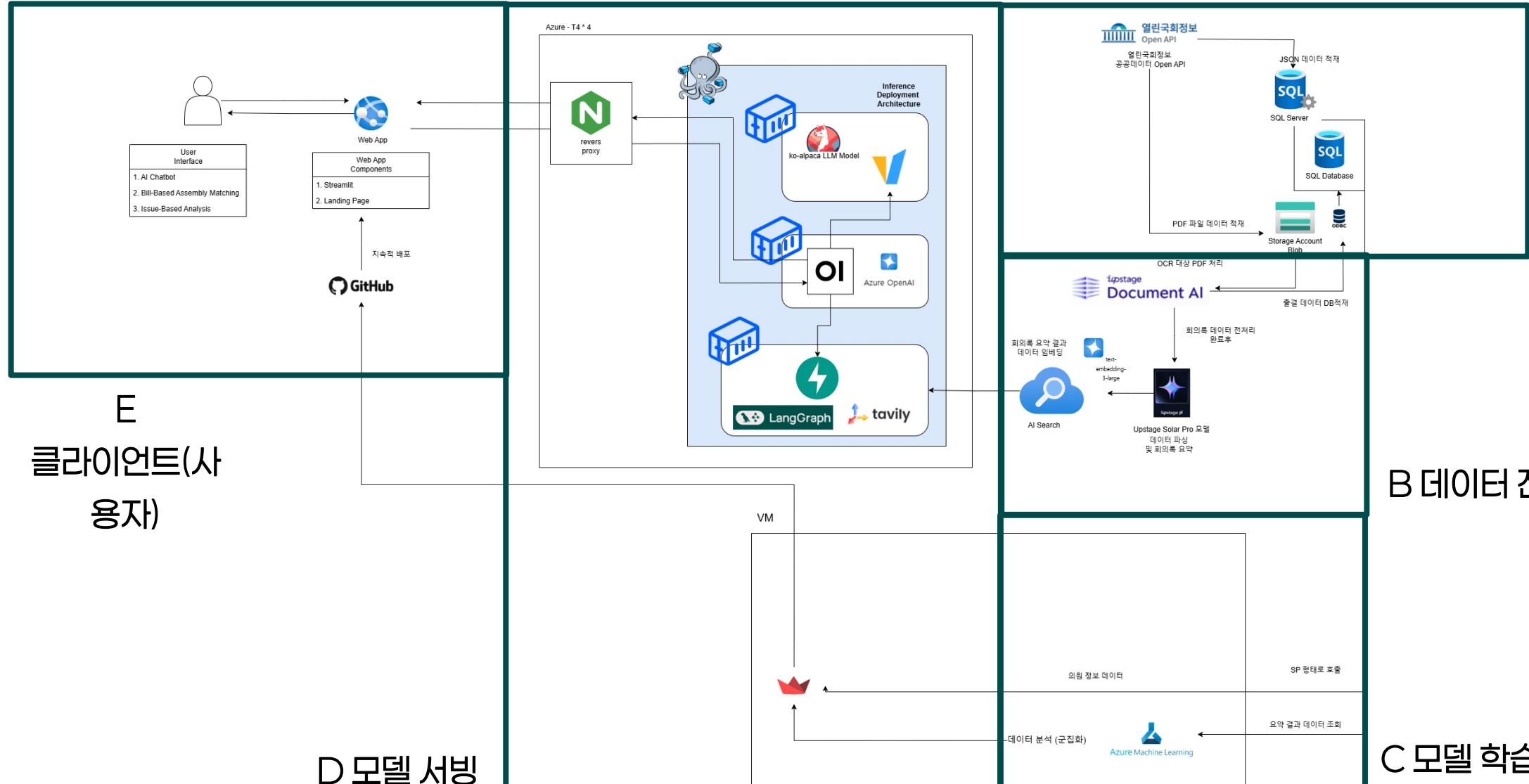


5-2. Architecture



5-2. Architecture

A 데이터 수집

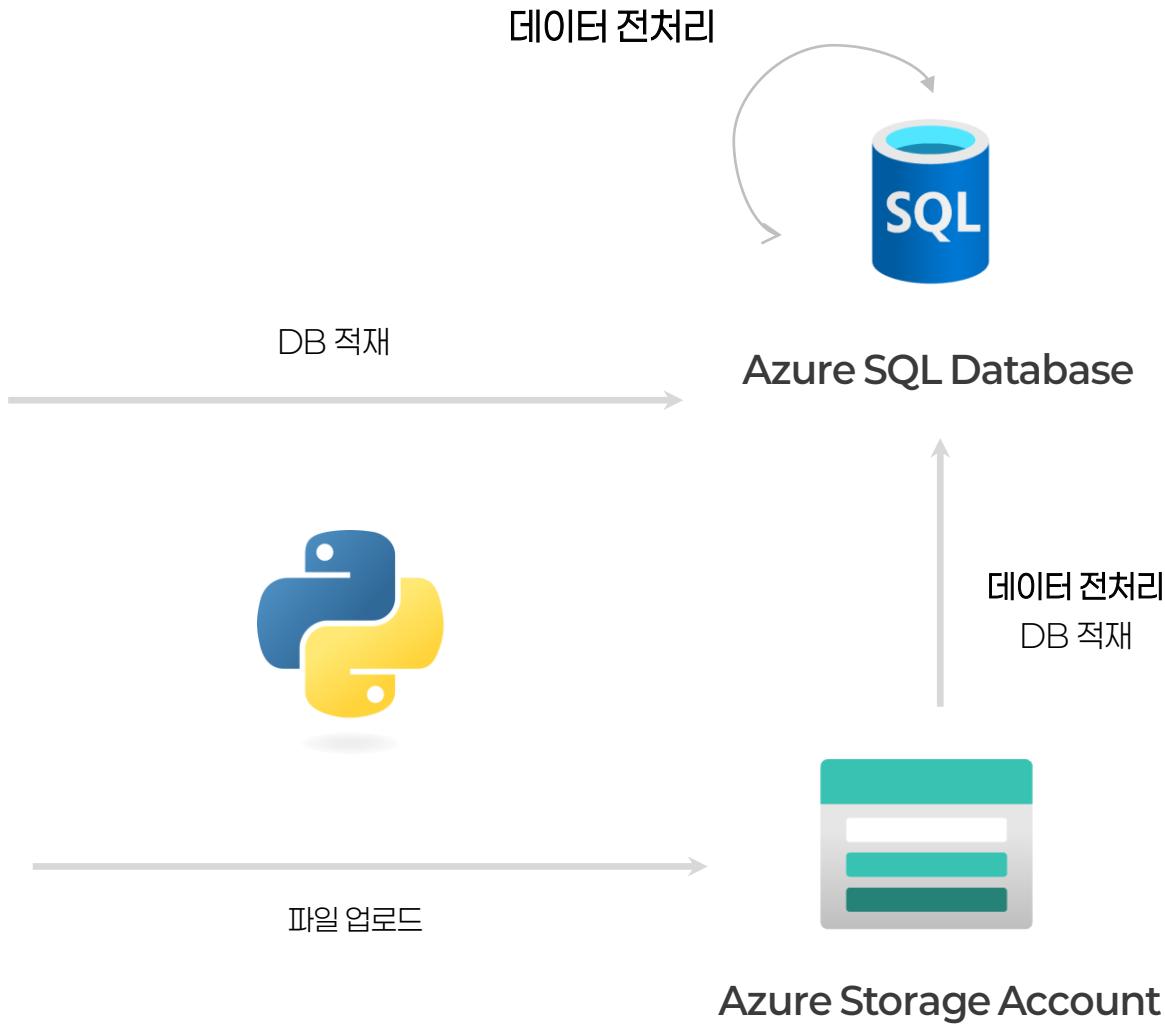


5-3. A 데이터 수집

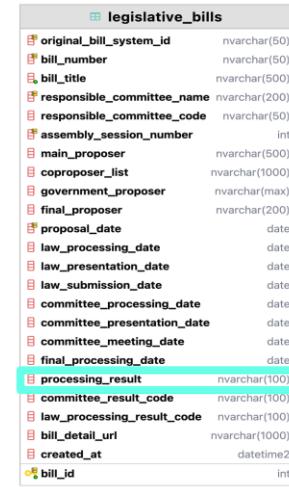
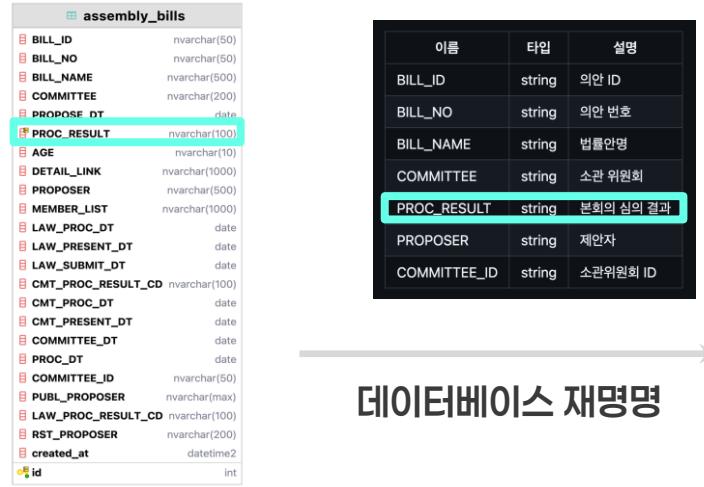


국회의원 본회의 표결 데이터 (약 50만건)
국회의원 발의법률안 (약 9만건)
국회의원 정보 통합 데이터 (약 3000건)
역대 국회의원 현황 (약 3000건)
국회의원 인적사항 (약 300건)

의안 회의록 PDF (약 1만 5천개)
국회의원 출석 PDF (약 130개)



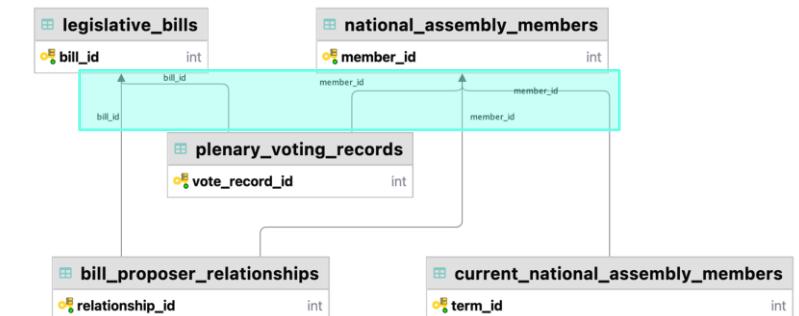
5-3. B 데이터 전처리(Tabular data)



데이터베이스 재명명



데이터베이스 정규화
(중복 최소화 및 구조 재정의)



5-3. B 데이터 전처리(이미지 > OCR)

의안 회의록 PDF (약 1만 5천개 : PDF 1개 약 30 ~ 150Page)



국회의원 출석 PDF (약 130개)

의원명	법제사법위원회					
	제424회		제425회		회의일수	출석
	(16/10) 09월 01 일	(16/10) 09월 02 일	(16/10) 09월 03 일	(17/10) 09월 04 일		
정희래(鄭熙來)	출석	출석	출석	4	4	0 0 0 0
유기별(劉基別)	출석	출석	출석	4	4	0 0 0 0
박병재(朴炳載)	출석	출석	출석	4	4	0 0 0 0
곽기철(郭基哲)	출석	출석	출석	4	4	0 0 0 0
박종현(朴宗鉉)	출석	출석	출석	4	4	0 0 0 0
김재현(金在顯)	검색	출석	출석	4	3	1 0 0 0
송석준(宋錫俊)	출석	출석	검색	4	3	1 0 0 0
장동혁(張東赫)	출석	출석	검색	4	2	2 0 0 0
조재숙(趙在淑)	검색	출석	출석	4	3	1 0 0 0
주진우(朱晉祐)	출석	출석	출석	4	4	0 0 0 0
김기표(金基彪)	출석	출석	출석	4	4	0 0 0 0
김용민(金容民)	출석	출석	출석	4	4	0 0 0 0
박근학(朴根學)	출석	출석	출석	4	4	0 0 0 0
박근원(朴根元)	출석	출석	출석	4	4	0 0 0 0
박희승(朴熙勝)	출석	출석	출석	4	4	0 0 0 0
서경교(徐景敎)	출석	출석	출석	4	4	0 0 0 0
이성윤(李成潤)	출석	출석	출석	4	4	0 0 0 0
장은태(張恩泰)	출석	출석	출석	4	4	0 0 0 0
박은정(朴銀貞)	출석	출석	출석	4	4	0 0 0 0

인덱스 스키마 정의: document_id(Key), content 등 메타데이터 필드
 Vector Search 설정: embedding 필드(3072차원), HNSW 알고리즘, cosine 유사도
 문서 업로드: 원본 데이터와 임베딩 벡터를 함께 인덱스에 구축



AI 검색



OCR				
구분	항목	활용 전	활용 후	비고
성공	정상 추출	117건	135건	2건 외 추출 완료
예외	OCR API 실패	5건	0건	해결 완료
	텍스트 오류	10건	0건	해결 완료
	패턴 매칭 실패	5건	2건	2건 잔여



Azure OpenAI

text-embedding-3-large

- ✓ 컨텍스트 보강 텍스트를 입력값으로 고차원 백터생성
- ✓ Batch 방식으로 데이터 핸들링



SQL 데이터베이스

5-3. C 군집화 모델 학습: 의원들의 표결활동을 바탕으로 한 군집화

1. 데이터 전처리

	법안1	법안2	법안n	
의원A	찬성	찬성	...	반대
의원B	반대	반대	...	찬성

의원Z	찬성	불참	...	찬성

- 의원 X 표결 활동 매트릭스 생성
- 각 법안에 대한 투표결과를 **One-Hot Encoding** 변환
[찬성: 1, 반대: 2, 기권: 3, 불참: 4]

2. 군집화 모델 학습

군집 수 2~8개로 모델 학습 후 비교

군집화 알고리즘:

- K-Means
- DBSCAN
- Gaussian Mixture Model (GMM)
- Agglomerative Clustering

3. 모델 평가

- 각 알고리즘 별 Silhouette Score 계산
- 가장 높은 점수를 기록한 모델 선택

4. 결과 시각화

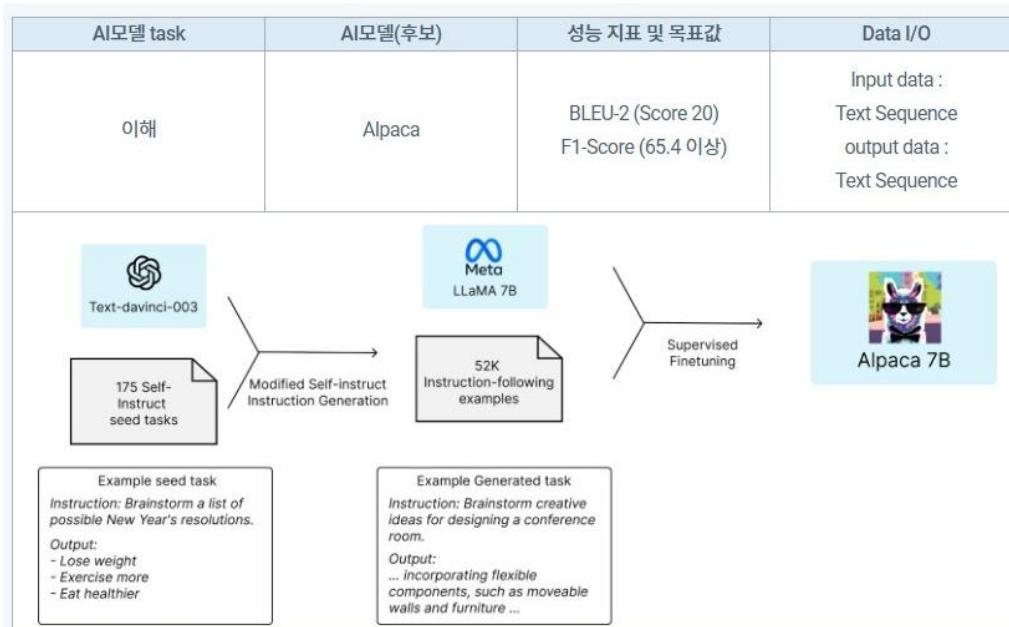
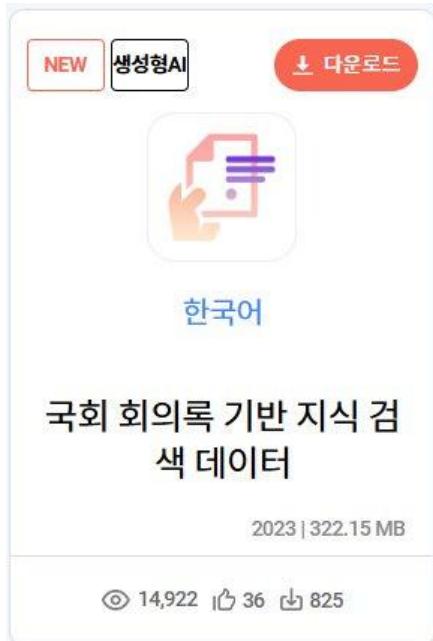
- 고차원 군집 결과를 PCA를 통해 2차원으로 축소
- 클러스터 별 성격 요약

5-3. D 모델학습 - LLM

모델 학습 진행 배경

- 1) Domain Specific 한 LLM 모델의 필요성 : 성능 저하 - 정치적인 질문의 경우 Block
- 2) AI-Hub에서 학습 데이터 획득

Preprocess > data split > train > inference



ko-Llama



- 📁 특별위원회
- 📁 소위원회
- 📁 예산결산특별위원회
- 📁 국정감사
- 📁 본회의

※ 15~20대 국회 데이터 기준 학습

5-3. D 모델학습 - LLM

모델 학습

KoAlpaca-Polyglot-5.8B (v1.1b)	
항목	사양
기반 모델	Polyglot-ko 5.8B
파라미터 수	
파라미터 수	6B (60억)
학습 하이퍼파라미터	
Learning Rate	5e-05
Train Batch Size	2
Optimizer	Adam (betas=(0.9, 0.999), epsilon=1e-08)
LR Scheduler	linear
Epochs	2.0
Mixed Precision	Native AMP
프레임워크 버전	
Transformers	4.29.0.dev0
Pytorch	2.0.0+cu117
Datasets	2.10.1
Tokenizers	0.13.2

Learing Rate 5e-05
Train Batch Size 2
Optimizer Adam
Epochs 2.0

성능의 이슈 문제로 > Azure Open AI로 대체

단, 학습 시킨 모델 webUI에서 확인 가능

koalpaca-pretrain +

모델 검색

전체 koalpaca-pretrain (app/model)

koalpaca-pretrain ⚡

gpt-35-turbo ⚡

gpt-40-mini ⚡

임시 재생

안녕하세요 최초 장애인 관련 법안에 대한 별의를 한사람은 누구인가요?

개인 관련 법안에 관련해서 별의한 것이 있으며, 실명화를 주장한 사람이 누구인가요?

날짜 오전 1:14

궁금해요.

법률적 주제

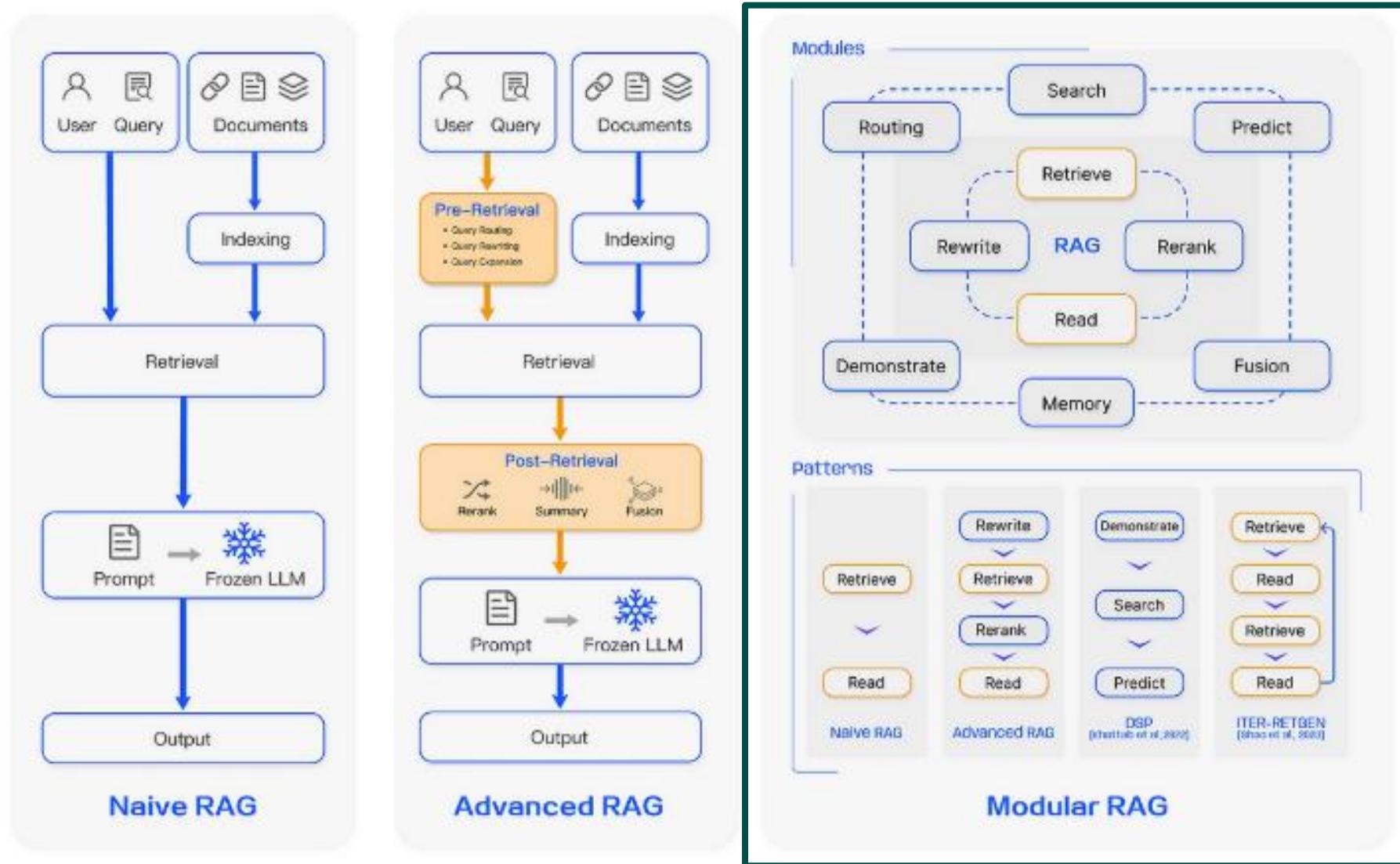
를 포함하고 있습니다. 예를 들어, 기초연금에 대한 공정성 문제나 기초보증금에서 발생할 수 있는 불이익 문제와 같은 주제들이 있습니다.

- 법률 제정 및 수정: 변호사들은 법률 제정 및 수정 과정에서 의견을 제시하거나 자문을 제공합니다. 이를 통해 법의 공정성을 높이고 법적 사각지대를 해소하고자 합니다.
- 사례 연구: 특정 사건이나 사례를 바탕으로 법률적 해석이나 판단을 내립니다. 이는 법원이 판단할 때 중요한 요소가 될 수 있습니다.
- 공청회 및 토론회: 다양한 법률적인 문제에 대해 공청회나 토론회에 참가하여 자신의 의견을 펼치고 대중에게 알리는 역할도 수행합니다.
- 소송 및 상담: 실제 소송을 수행하거나 클라이언트에게 법률 상담을 통해 그들의 권리를 보호하는 작업도 포함됩니다.

이와 같은 활동을 통해 변호사들은 사회의 법률 문제를 해결하고, 공정하고 정의로운 법적 환경을 조성하는 데 기여하고 있습니다. 변호사들이 발휘하는 내용이나 역할에 대해 더 궁금한 점이 있으시면 말씀해 주세요! 안녕하세요! 변호사들이 발휘한 내용에 대해 궁금하신 점을 말씀해 주셔서 감사합니다. 변호사들은 다양한 법률적 주제를 다루며 여러 가지 역할을 수행합니다. 여기에는 다음과 같은 내용이 포함될 수 있습니다.

- 법률 자문 및 상담: 개인이나 기업에게 법률 문제에 대한 조언을 제공하고, 복잡한 법률 문서를 이해하기 쉽게 설명합니다.
- 소송 대리: 법원에서 고객을 대리하여 소송을 진행하고, 변론을 통해 고객의 이익을 보호합니다.
- 계약 및 협상: 다양한 계약서를 작성하고 검토하며, 계약의 조건을 협상하여 고객의 권리와 최대한 보장합니다.
- 법률 연구 및 해석: 새로운 법률이나 판례를 연구하여 고객에게 관련 정보를 제공하고, 법적 쟁점을 명확히 합니다.
- 공공 정책 및 입법 활동: 특정 사회적 이슈에 대해 의견을 전달하고, 법률 개정 또는 제정을 위한 활동에 참여합니다.
- 교육 및 공공 소통: 일반 대중에게 법률 지식을 전달하고, 권리와 의무에 대한 인식을 높이는 작업도 수행합니다.

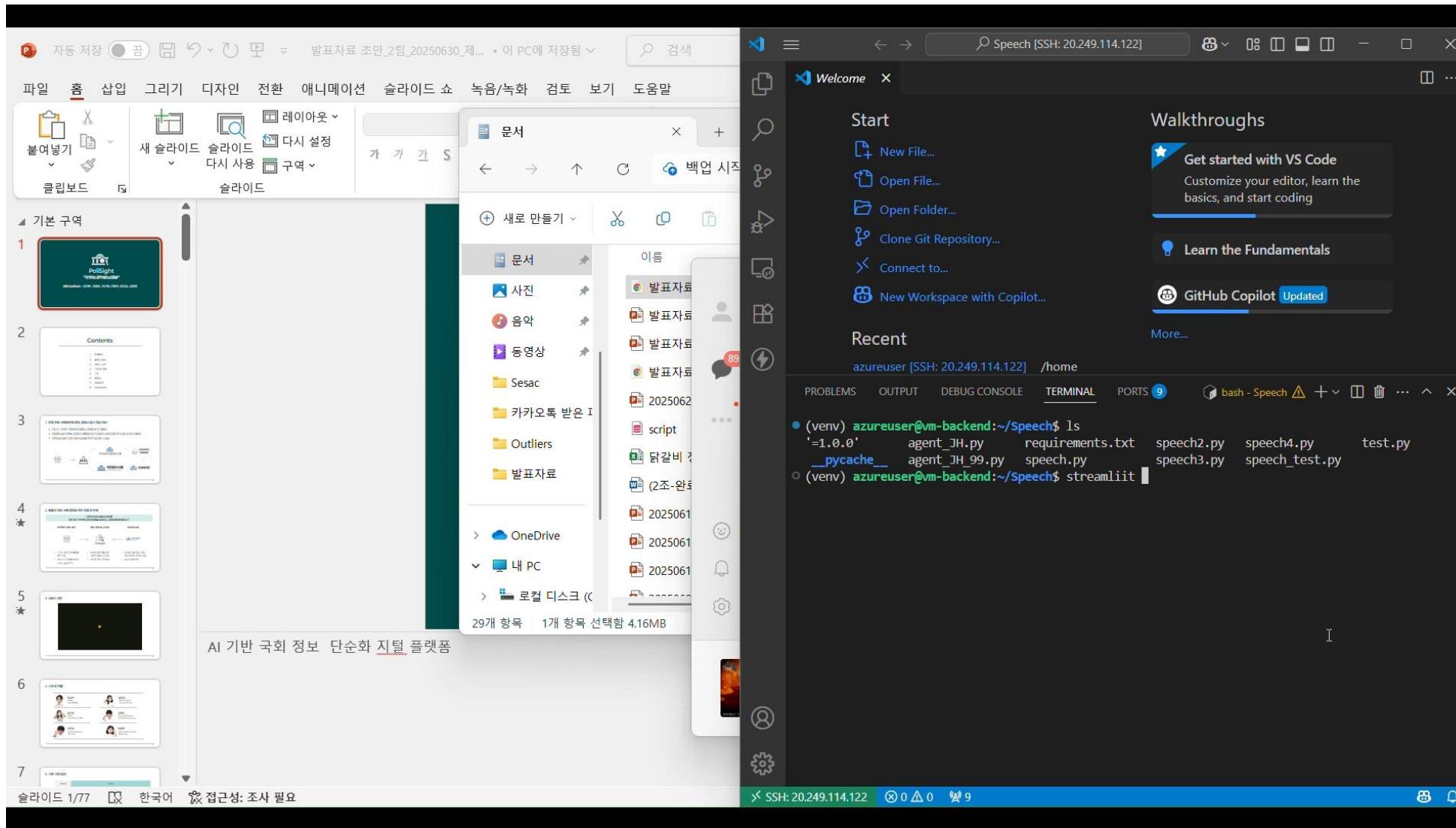
5-3. D 모델학습 - LLM



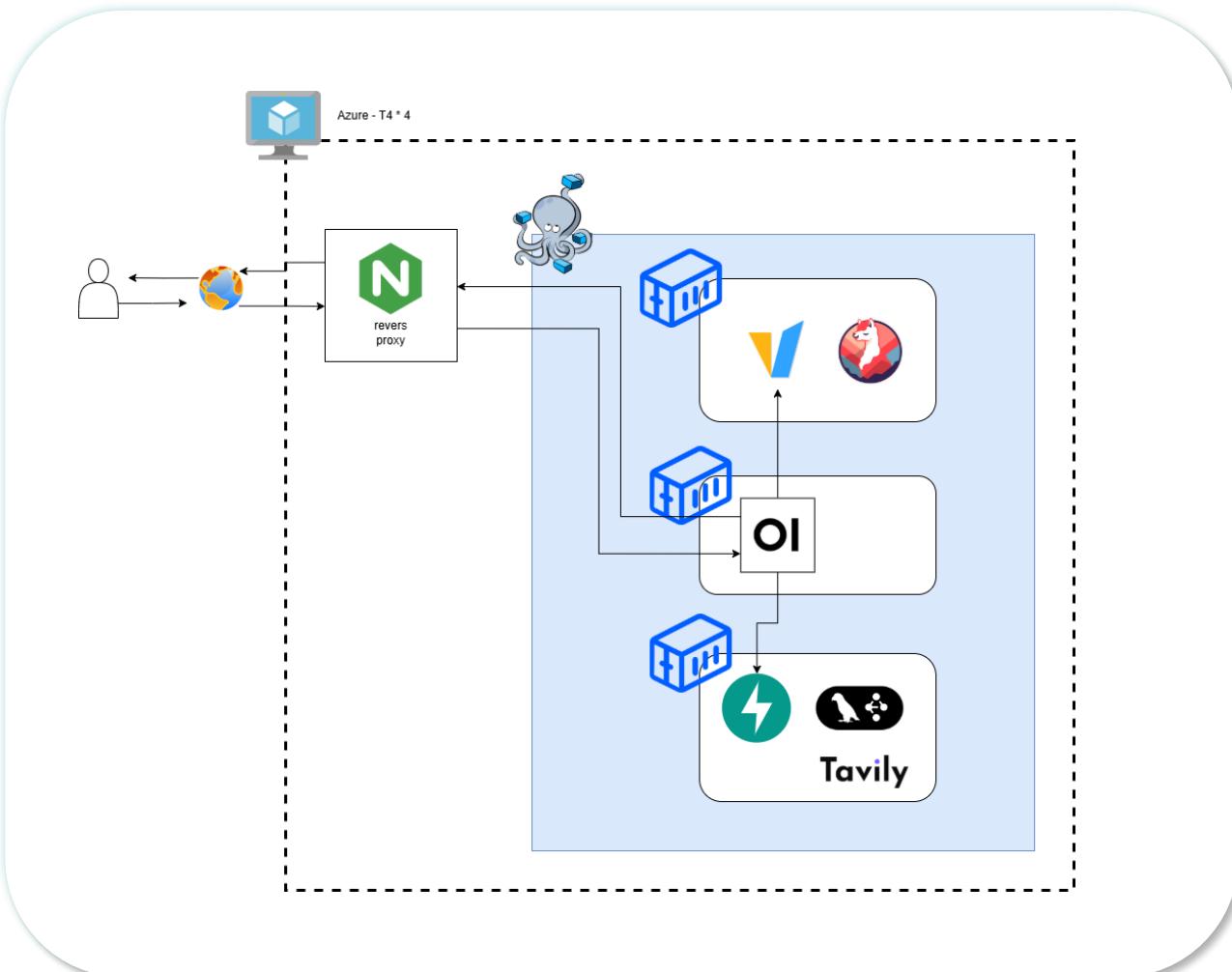
5-3. D 모델학습 - Modular RAG



5-3. D 모델학습 - Modular RAG: 분기 조건에 따른 실행 경로의 실제 작동 여부



5-3. D 서빙 아키텍처



Nginx rever porxy
open webui로 연결하고
compose로 컨테이너 관리

학습 시킨 모델 인퍼런스

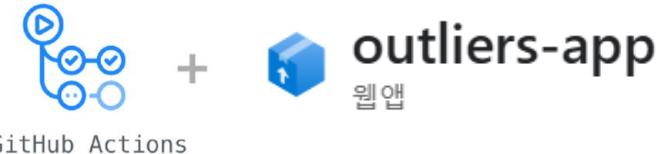
Backend 필요 = VLLM 으로 배포

Modular RAG Tool calling 용 Fast API

이후 Open WebUI 연결
Azure Service 이용하여
STT/TTS 세팅

5-3. E 클라이언트(사용자) 구성하며 고민했던 것들

Azure App과 Github Action 연동
= 지속적인 배포관리



GitHub Actions

The screenshot shows the Azure DevOps Pipeline Center for the "outliers-app" project. On the left, there's a sidebar with various service links like Microsoft Defender, Event Grid, Log Analytics, and Container Registry. The main area displays a table of deployment logs:

날짜	ID	작업 로그	작성자	상태	설명
2025. 6. 26., 06:01:33 PM	e1cb463	업 로그	N/A	성공	OneDeploy
2025. 6. 26., 03:52:00 PM	f5f3883	업 로그	N/A	성공	OneDeploy
2025. 6. 26., 03:48:11 PM	149d84a	빌드/배포 로그	chloecho22	성공	plot
2025. 6. 26., 02:40:50 PM	abcd802	업 로그	N/A	성공	OneDeploy
2025. 6. 26., 02:37:17 PM	4741681	빌드/배포 로그	chloecho22	성공	test

Below the table, a detailed view of a specific deployment is shown, including steps like "Merge pull request #1 from Outlier-sesac/JA #22" and "Deploy to Azure Web App".

대화 인터페이스를
Streamlit > WebUI로

The screenshot shows a Streamlit application titled "Agent 연동 음성 대화 AI" (Agent Integration Voice Conversation AI). It features a sidebar with "시스템 상태" (System Status) and "작업 목록" (Job List), and a main area with a text input field and a "작업 목록" (Job List) section. The right side of the screen shows a "gpt-4o-mini" chat interface with a message history:

- 오늘 어떻게 도와드릴까요?
- + ⚡ 1 코드 인터프리터
- 제안
- Give me ideas
for what to do with my kids' art
- Overcome procrastination
give me tips
- Tell me a fun fact
about the Roman Empire

5-4. E 클라이언트(사용자) 구성 시 고려한 핵심 설계 요소



The screenshot shows a user interface for an AI application. At the top, there is a search bar with the placeholder text "검색어를 입력하세요". Below the search bar, a button labeled "인증하기" (Authenticate) is visible. The main area features a profile picture of a person with short brown hair and the text "gpt-4o-mini". A message input field contains the Korean text "오늘 어떻게 도와드릴까요?". Below the input field, there are three suggested prompts: "Give me ideas for what to do with my kids' art", "Overcome procrastination give me tips", and "Tell me a fun fact about the Roman Empire". On the right side of the interface, there are icons for microphone and a circular profile picture.

기대 효과

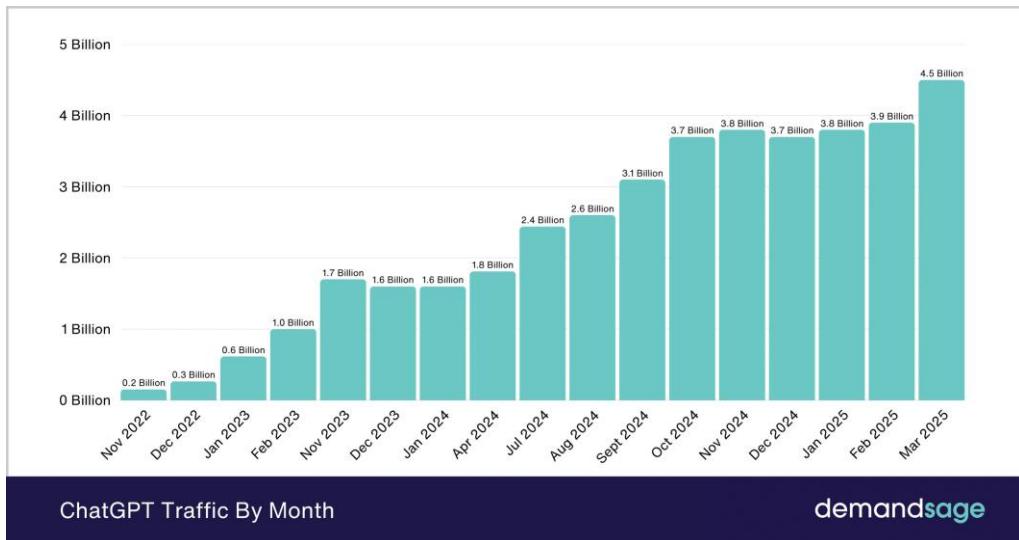
소외된 국민 없는, 모든 국민의 적극적인 정치 참여 지원



서비스 확장성

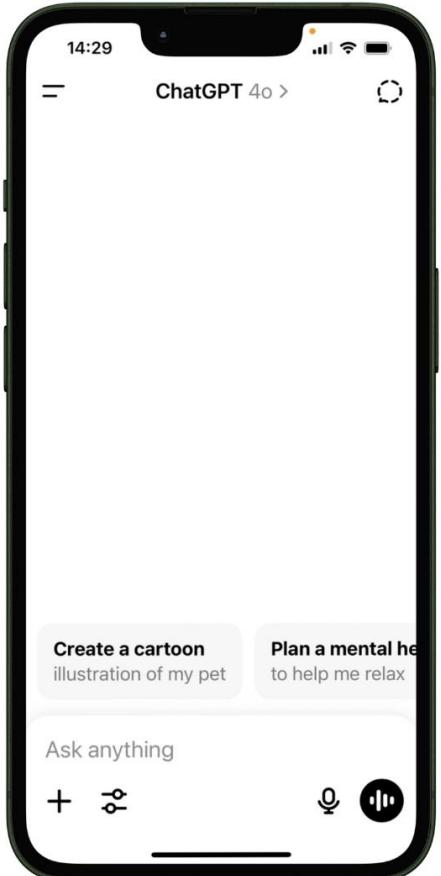
왜 음성 인터페이스가 중요할까요?

ChatGPT MAU (월간 활성 사용자)
: 2025년 5월 기준 약 45억명



Age Group	Share of ChatGPT Users
18-24 years	24.27%
25-34 years	30.58%
35-44 years	19.02%
45-54 years	12.89%
55-64 years	8.10%
Over 65 years	5.15%

서비스 확장성



- AI 서비스의 표준,
ChatGPT의 음성 인터페이스 기능
- 접근성
이미 AI 서비스를 경험한 고객을 위한 Open WebUI 도입
- 포용성
모두를 위한 STT, TTS 기능으로 AI UX 전략 실현

시장 가능성 - 잠재 수요층

< 부문별·계층별 디지털정보화수준(23년 : ■ / '24년 : ■) >

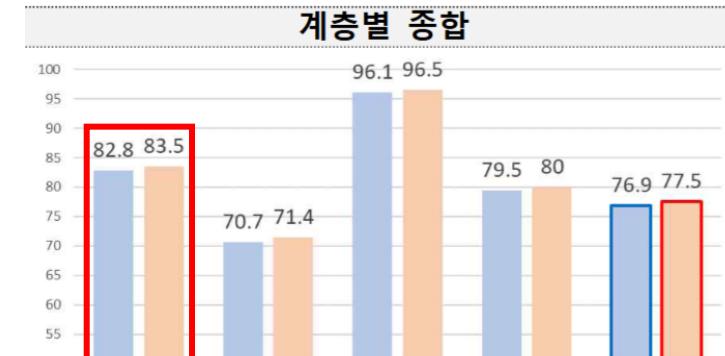
지난 20대 대통령 선거(대선)에서 투표 여부

(단위: %, 명)

구분	지체 장애	뇌병변 장애	시각 장애	청각 장애	언어 장애	지적 장애	자폐성 장애	정신 장애	신장 장애	뇌전증 장애	전체
예	88.6	63.1	84.6	89.4	79.4	55.1	53.7	62.1	84.6	75.0	82.1
아니오	11.4	36.9	15.4	10.6	20.6	44.9	46.3	37.9	15.4	25.0	17.9
계	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
전국추정수	1,160,334	244,503	252,094	418,039	17,660	160,246	16,215	103,780	107,171	6,812	2,538,516

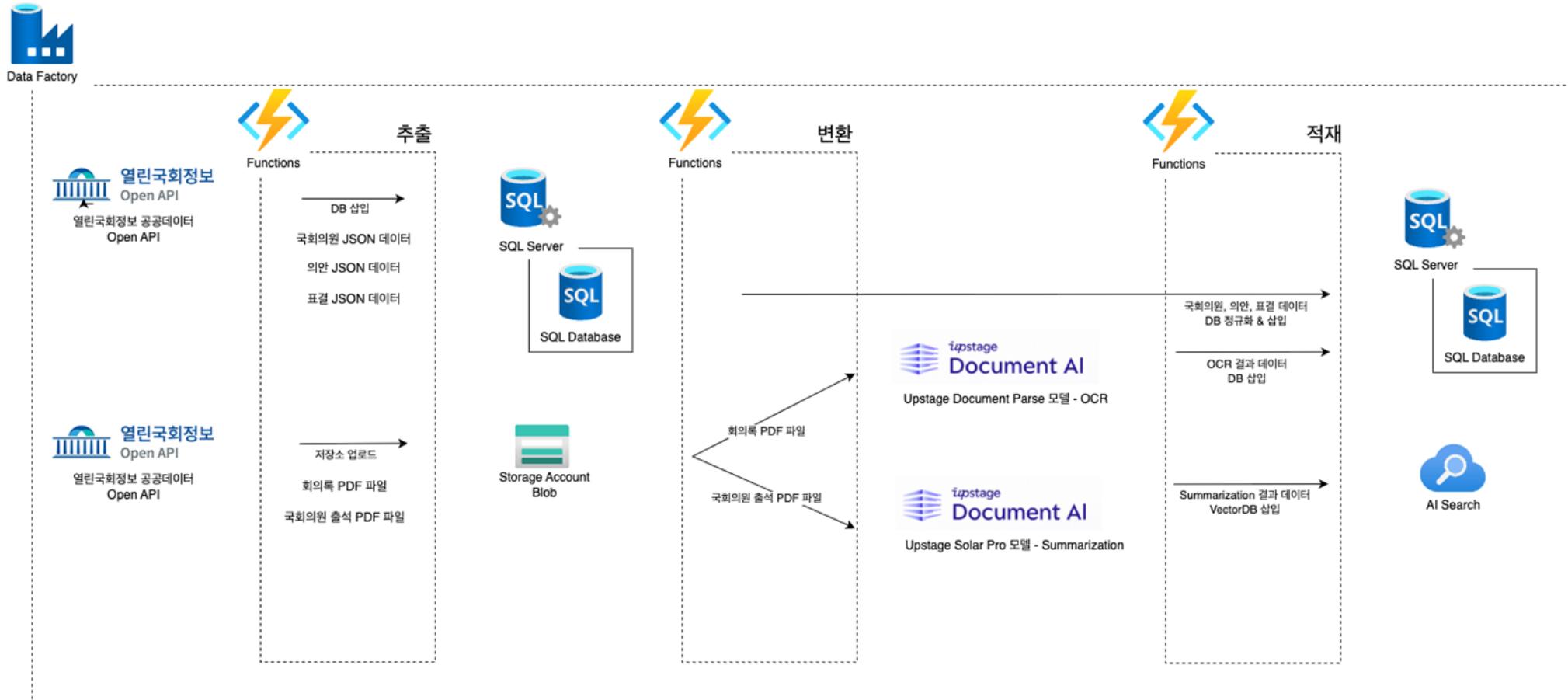
선거에 대한 정보를 얻는데 가장 도움이 되는 매체나 정보원

구분	지체 장애	뇌병변 장애	시각 장애	청각 장애	언어 장애	지적 장애	자폐성 장애	정신 장애	신장 장애	뇌전증 장애	전체
TV신문 등 언론보도와 기사	37.7	34.7	35.3	35.9	36.1	28.6	26.8	37.2	37.5	35.0	36.1
TV 대담, 토크회 등 방송연설	31.1	31.7	29.8	25.9	31.0	10.9	7.0	18.2	35.4	25.0	28.5
가족·친구·이웃 등 주변사람과의 대화	17.5	21.7	21.9	25.9	17.6	47.5	45.9	28.1	12.2	24.4	22.0
후보자의 홍보물 (선거공보, 현수막 등)	3.7	2.6	3.3	4.9	4.3	4.6	4.4	5.4	3.8	4.5	3.9
선거관리위원회의 우편물 (후보자, 공약 정보)	5.2	3.7	5.2	5.0	4.3	3.6	7.4	5.1	3.8	2.0	4.8
인터넷, SNS를 통한 정보	4.9	5.0	4.5	2.2	6.7	3.2	6.0	4.9	7.3	9.2	4.4



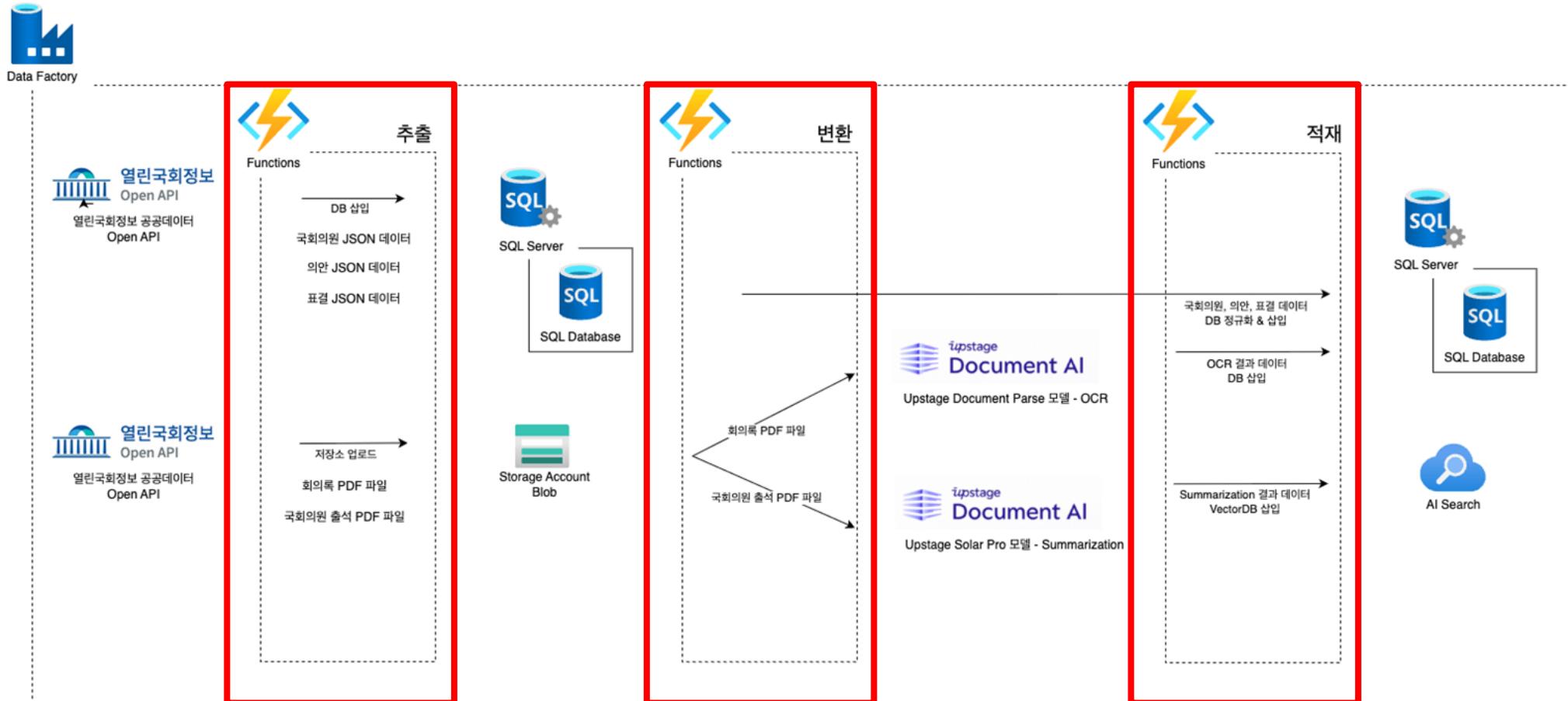
서비스 확장성 – 기술(데이터 파이프라인 구축)

- 목적: 지속적인 운영을 위해 데이터 수집 자동화



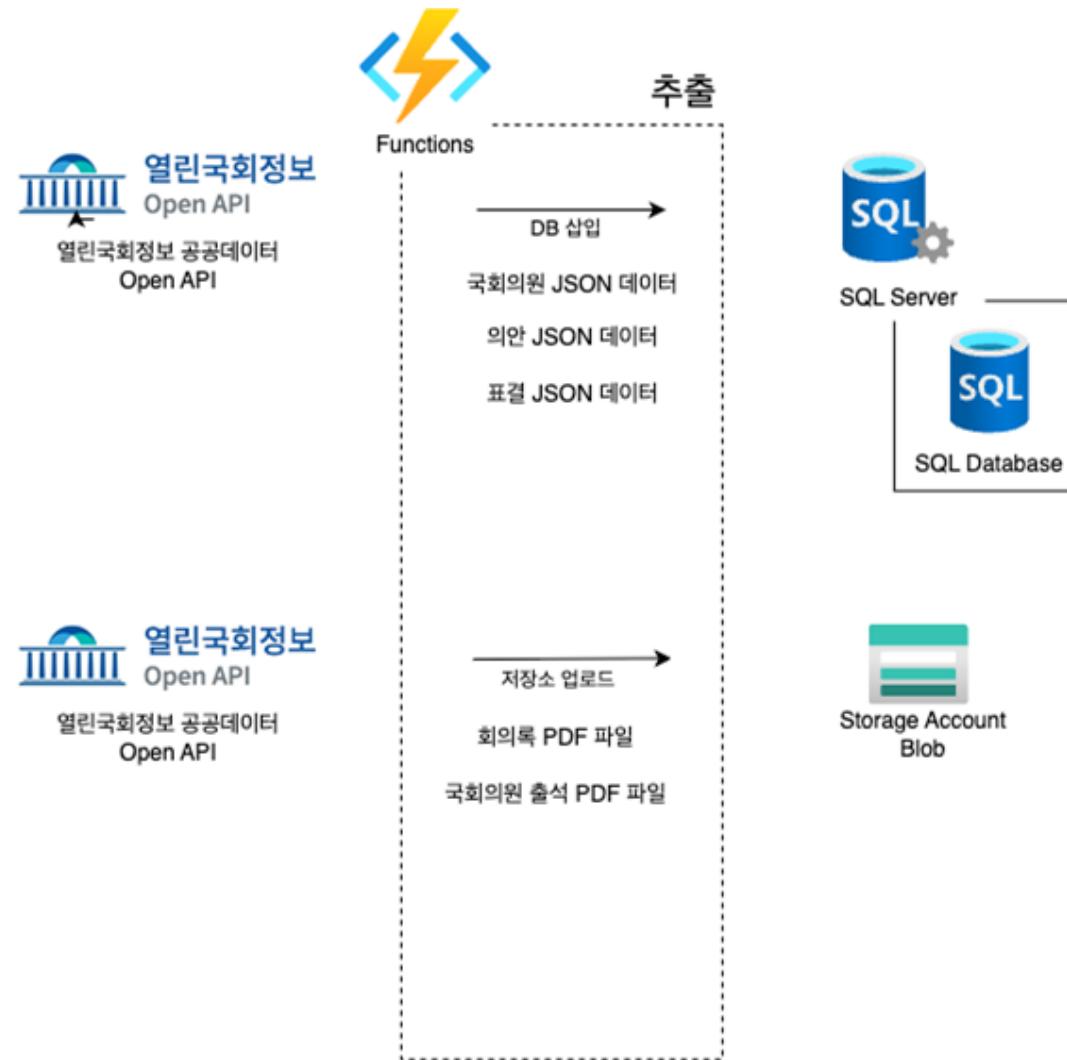
서비스 확장성 – 기술(데이터 파이프라인 구축)

- 목적: 지속적인 운영을 위해 데이터 수집 자동화



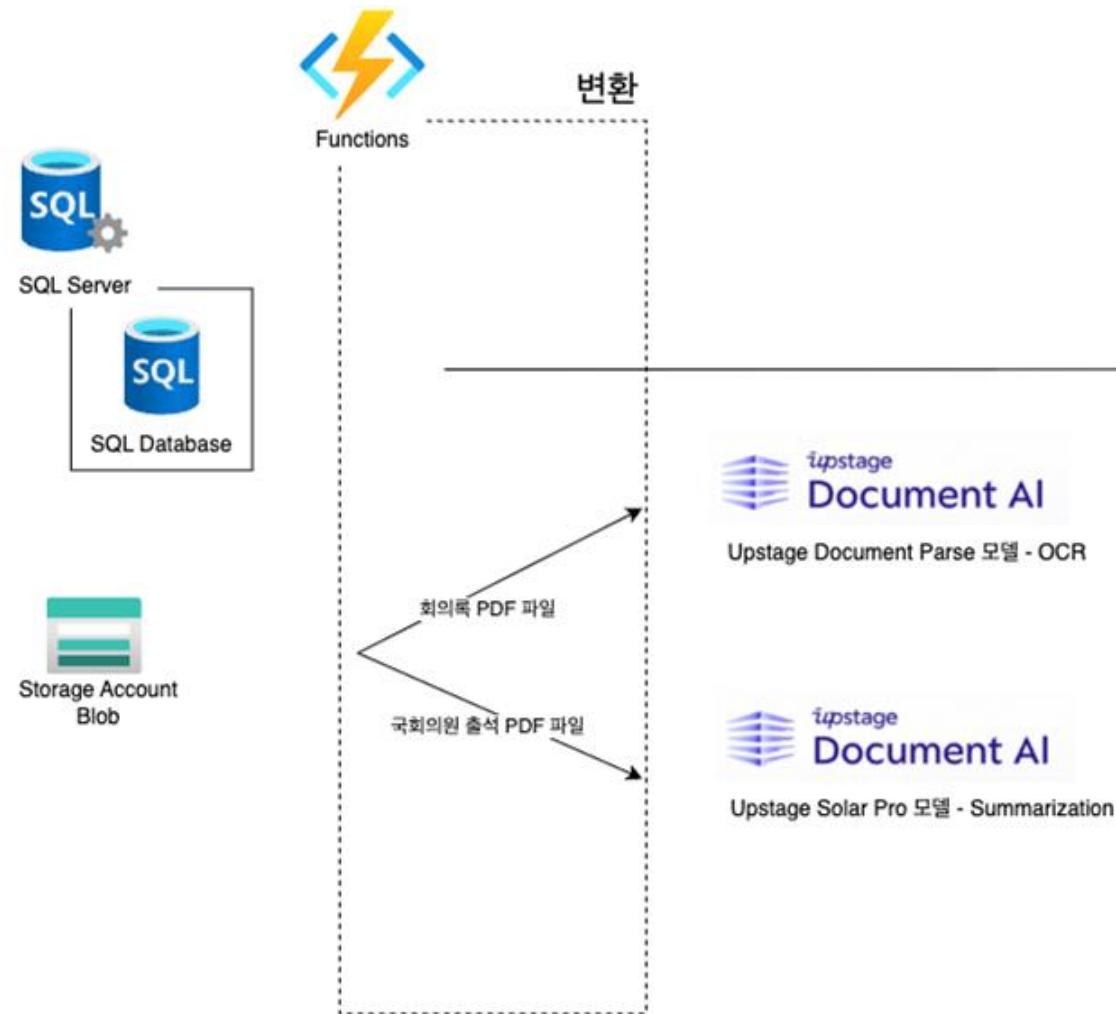
서비스 확장성 – 기술(데이터 파이프라인 구축)

- 추출 (Extract)



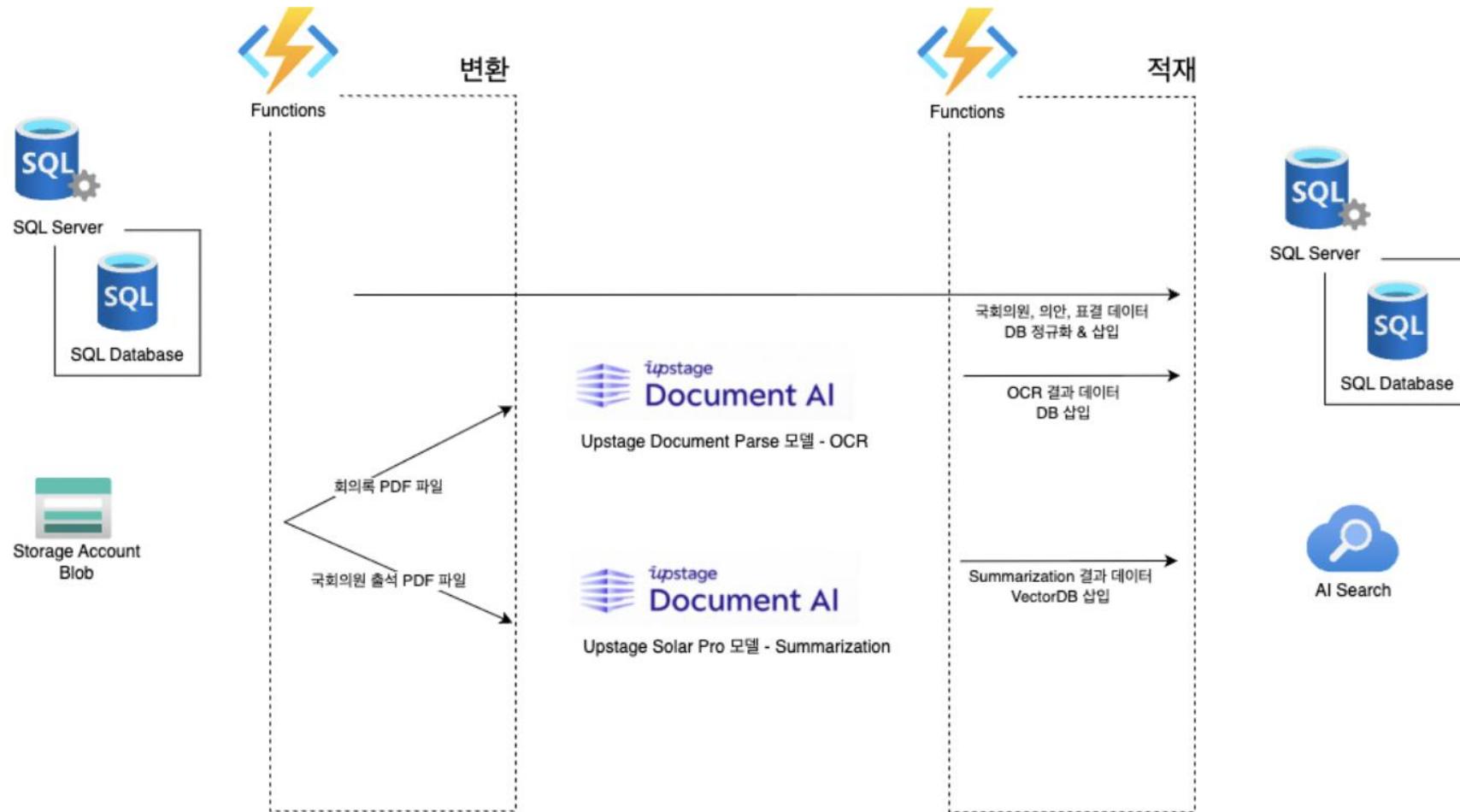
서비스 확장성 – 기술(데이터 파이프라인 구축)

- 변환 (Transform)



서비스 확장성 – 기술(데이터 파이프라인 구축)

- 적재 (Load)



인공지능의 윤리원칙

포용성

비단 일반 인원 뿐만 아니라, 청각/시각 장애인을 또한 사용할 수 있도록 aria-label 기입

투명성

국회 API/PDF만 활용, LangGraph 기반 RAG 시스템 답변 생성시 근거자료 제공

책임성

운영 정책을 기재하여 정보 소재 및 면책 사항 명시

신뢰성, 안정성

Azure 클라우드 기반 안정적 인프라 구성. AZURE OPEN AI 의 경우 데이터 미학습

개인정보보호

개인정보 필수 수집 사항 배제

보안

MS SQL Server SP 구성을 통한 안정적인 시스템 제공 및 SQL Injection 보안 공격 방지

자체평가 - 잘한 점

사용자

- 질문을 받고 신뢰도 높은 국회 공식 기록과 최신 웹 정보를 종합하여 맞춤형 답변을 제공
- 음성 답변을 통해 포용성 높은 사용자 경험을 제공

기술

- LangGraph를 활용해 질문의 의도를 분석하고 최적의 검색 전략(내부/외부/혼합)을 동적으로 결정
- 병렬 처리와 데이터 캐싱으로 응답 속도를 높임

비용

- Azure 클라우드 서비스에서 약 10개 이상의 리소스를 사용하면서도 철저한 사전 비용 계산과 Azure 비용 경고 기능을 활용하여 정해진 예산 내 MVP 개발 완료

자체평가 - 보완할 점

사용자

최초 검색 결과가 모호한 경우, 사용자에게 다시 질문하거나 선택지를 제공하여
의도를 명확히 파악하고 정확한 답변을 생성하도록 보완
한 답변에 여러 정보 출처("[국회 회의록]", "[웹 검색]")를 명확히 표기하여
신뢰성을 높이고, 사용자가 추가 정보를 확인할 수 있게 안내하도록 보완

기술

Azure AI Search의 하이브리드 검색(벡터 검색 + BM25 키워드 검색) 기능을 도입하여
의미적 유사성과 키워드 정확도를 모두 잡아 검색 품질을 향상하도록 보완
검색 결과에서 여러 출처의 정보들을 크로스-인코더 기반의 리랭커(Re-ranker)로
재정렬하여 최종 답변 생성 LLM에게 가장 관련성 높은 정보를 우선적으로 제공하도록 보완

인프라

Azure 클라우드 포털의 GUI 기반 리소스 관리를
IaC(Infrastructure as Code) 도구를 통해 코드 기반의 선언적 관리 방법으로 전환하여
인프라 관리의 일관성과 유지보수성을 향상하도록 보완

감사합니다

구현한 서비스 직접해보기



- Outliers

Appendix - 스크립트 평가 프로젝트(실패 결과 공유)

AutoML

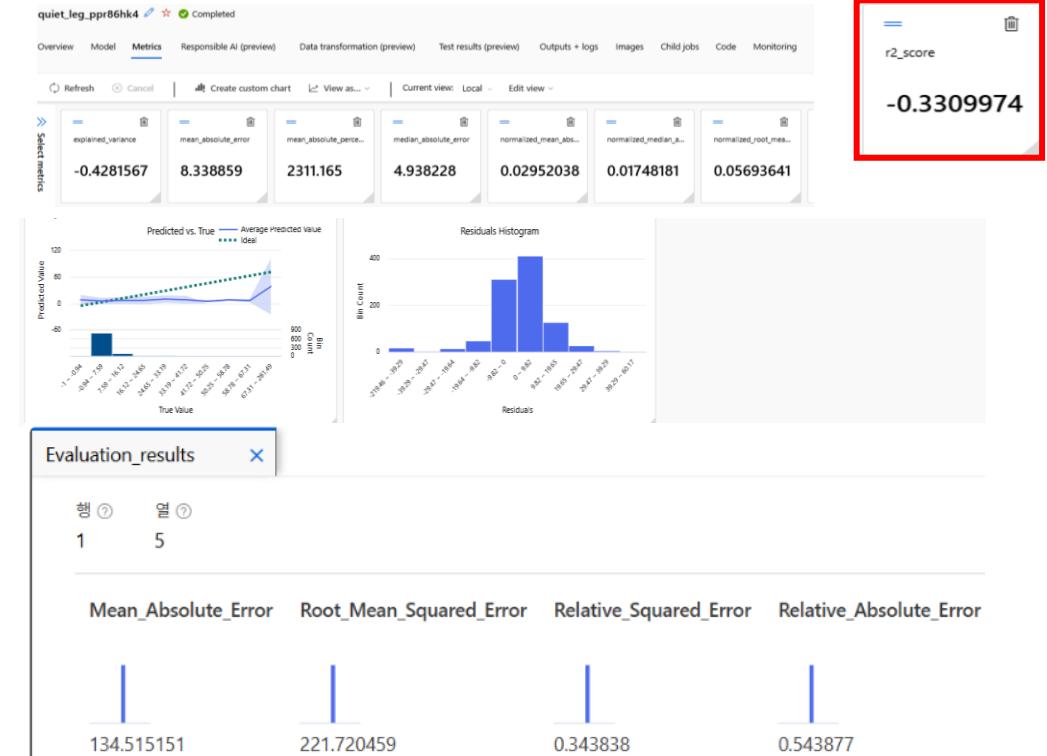
(MaxAbsScaler, LightGBM 설정)

script, roi(타깃) 이용 r2 score가 -로 진행 불가 확인

Result

Boosted Decision Tree Regression

- 특성: (script 제외)genres, release_year, runtime, vote_average, popularity, keywords, roi(타깃)
- 결과: MAE 134.5로 roi 평균(361) 대비 43% 오차, 상대 지표들(RSE, RAE)은 기준 모델 대비 낮은 오차값(34%와 54%) 확인



Appendix - 스크립트 평가 프로젝트(실패 결과 공유)

Cost

합계 (KRW) ⓘ

₩70,291

평균

<₩0.01 /일

결과 공유

서비스	제품	미터	요금 유형	합계 ⓘ
Virtual Machines	Virtual Machines NCast4 v3 Series	Virtual Machines NCast4 v3 Series - NC64asT4 v3 - KR Ce...	Usage	₩69,123
Storage	Premium SSD Managed Disks	Premium SSD Managed Disks - P10 - KR Central	Usage	₩550.15
Load Balancer				₩529.22

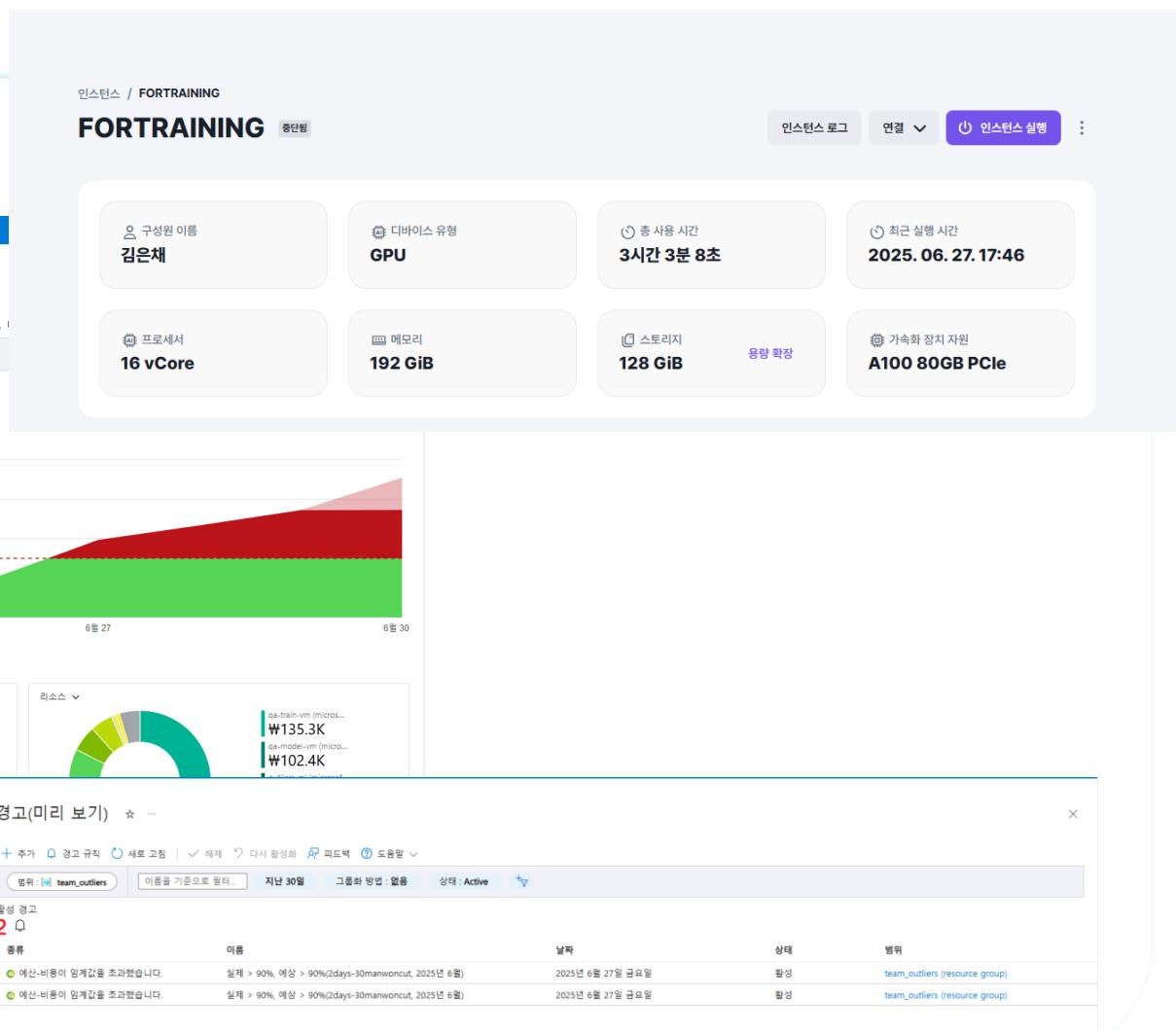
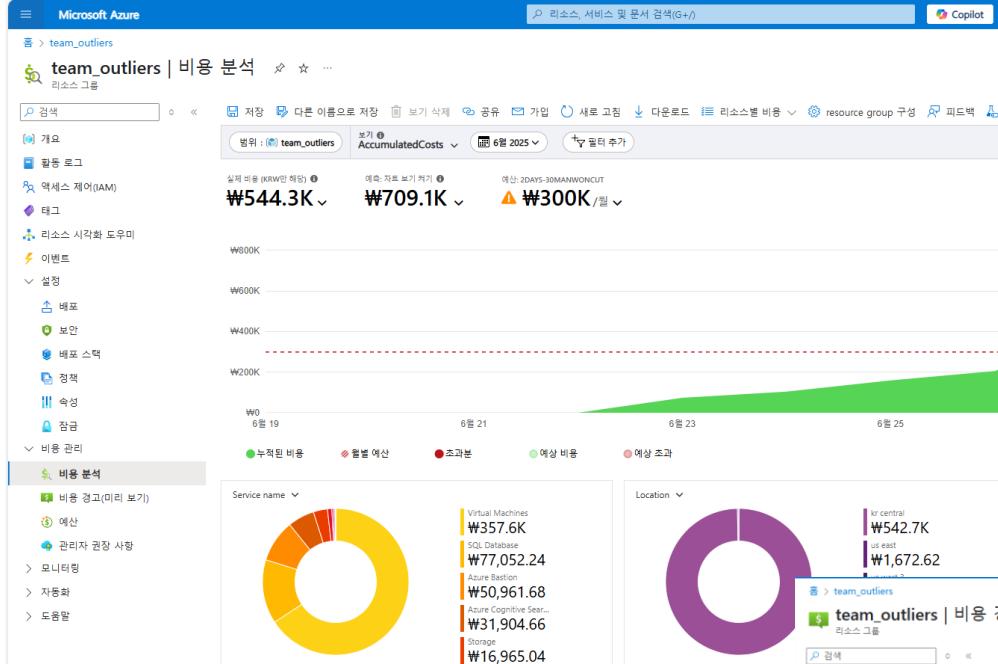
결과

- 간단한 NLP를 이용한 script 데이터는 기타 콘텐츠info 대비 낮은 예측값 확인
- 이는 원래 프로젝트 콘셉트에 맞지 않으며, 개선을 위해 오랜 NLP 작업 필요
- 짧은 기간 팀원들의 역량을 고루 이용하는 프로젝트로 변경 [ML 총 비용 70,291(일)]

Appendix - Cost

크레딧 W3,140

-06/30 Cost



Appendix - 왜 Bing Search 를 쓰지 않고 Travily API를 썼을까?

Azure Bing Search API 한계

1. EOS

2025년 8월 완전 폐쇄 예정

신규 사용자 등록 중단으로 사실상 사용 불가

2) 대체 서비스의 문제점

'Grounding with Bing Search'는

AI 요약 결과만 제공하여 원본 데이터 접근 불가

기존 대비 최대 483% 높은 비용 발생

응답 속도 저하로 실시간 검색에 제약

Travily API 선택 이유

1) RAG 최적화 기능

LLM이 직접 활용 가능한 구조화된 원본 검색

결과 제공

실시간 정보 검색에 특화된 AI 워크플로우 지원

2) 안정성과 확장성

서비스 중단 리스크 없는 지속 가능한 운영

합리적 비용 모델로 예측 가능한 확장성

LangChain, LangGraph와 생태계 통합

Appendix - 시장 가능성(경쟁사 서칭)



공공기관 주관 플랫폼

- 장점: 신뢰성 및 API 제공
- 한계: UX/UI의 경직성 및 이용 난이도

시민단체 운영 서비스

- 시민 참여 용이, 독립성
- 데이터 누락 및 편향, 운영 지속성의 위험

AI·빅데이터 기반 정보서비스

- 자연어 검색, 예측, 추천
- 포용성 관심 부족

Appendix - Assistive Tech 벤치마크

서비스 이름	기능	우리가 배운 점
Aria Label	스크린리더 등 보조기술에 음성 라벨 추가	청각적 접근성 방법론(실제 적용)
스크린 리더 (Screen Reader)	화면 내용을 음성/점자로 출력하는 소프트웨어	음성 인터페이스 설계 원칙과 사용성 개선
독서확대기	저시력인을 위한 화면 확대 및 OCR 기능	확대 가능 인터페이스 및 OCR 필요성 확인
VoiceOver (Apple)	macOS 내장 스크린리더 프로그램	음성 네비게이션과 키보드 조작 편의성 확인
Accessibility Insights	웹 접근성 테스트 브라우저 확장프로그램	접근성 테스트 및 검증 방법론
Be My Eyes	시각장애인과 자원봉사자 실시간 영상통화 매칭	차후 자발적 커뮤니티와 Assistive tech 결합 가능성
점자정보단말기	시각장애인용 노트북 (TTS, 점자셀 지원)	이후 점자정보 서비스 확대 가능성 탐진

Appendix - 시장 가능성 (관련 시장 규모 및 성장)

시장 분야	2024년 규모	2030년 전망	연평균 성장률	출처
시각장애인 보조기술	42억 달러	121억 달러	13.2%	'시각 장애 시장을 위한 보조 기술' (Kings 리서치)
TTS(음성합성)	44억 달러	79억 달러	10.2%	'TTS 시장 성장예측' (2025-2030년, Global Information)
음성인식 AI	184억 달러	517억 달러	23.0%	'음성인식 : 시장 점유율 분석, 산업 동향 및 통계, 성장 예측' (2025-2030년, Global Information)
대화형 AI	122억 달러	669억 달러	22.6%	'대화식 AI 시장 규모, 예측 2025-2032'(포츈 Business Insights)
디지털 접근성	12억 달러	43.2억 달러	13.8%	디지털 접근성 서비스 시장 규모, 점유율, 성장 (Business Insight Research)

Appendix - 목표 설정 - 예상 매출 및 비용 분석

목표 구분	세부 내용	예상 매출
단기 목표 (6개월)	시각장애인 사용자 1,000명 확보	프리미엄 서비스 월 50만원
단기 목표 (6개월)	국회 의원실/언론사 2-3곳 파트너십	B2B 컨설팅 월 300만원
중기 목표 (1-2년)	월간 활성 사용자 5,000명 달성	구독료 + API 수수료 월 2,000만원
중기 목표 (1-2년)	B2G 계약 통한 안정적 수익 모델 구축	정부 용역 + 기업 계약 월 5,000만원
장기 목표 (3-5년)	전체 시각장애인 10% 이용 (2.4만명)	다각화된 수익 모델 월 2억원
장기 목표 (3-5년)	정치 정보 접근성 표준 플랫폼 지위 확보	해외 진출 포함 월 5억원

항목	비용	세부 설명
누적 개발비 (10일)	60만원	Azure 클라우드, GPU VM 사용료
향후 개발비(6개월)	500만원	추가 기능 개발, 고도화
연간 서비스 운영비	Azure 클라우드 1,200만원	월 100만원 × 12개월
마케팅 비용	800만원	장애인 단체 협력, 광고
인건비 (6명)	3억 원	연봉 5,000만원 × 6명 (파트 타임)
총 연간 소요 비용	약 3억 2,500만원	손익분기점: 월 1,200만원 매출

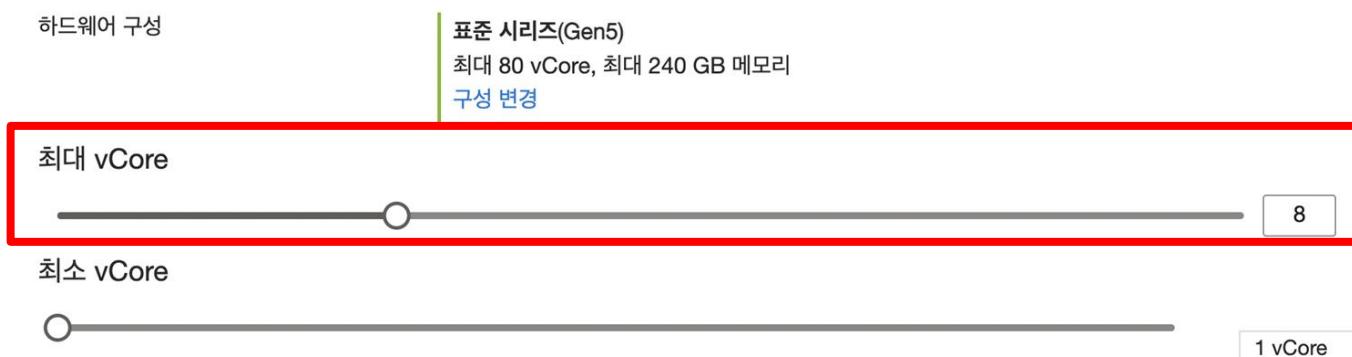
1. 트러블슈팅 - 데이터 수집

- 문제: 총 51만건의 데이터를 DB에 삽입하는 과정에서 너무 오랜 시간 소요
- 해결방법: Azure SQL Databases 최대 vCore 2개 -> 8개 조정
- 결과: 데이터 DB 삽입 시간 약 50~70% 감소

서비스는 Azure SQL 데이터베이스에서 단일 데이터베이스를 위한 컴퓨팅 계층으로, 워크로드 수요에 따라 컴퓨팅 자원을 자동으로 조정하고 초당 사용된 컴퓨팅 양에 대해 청구하는 방식입니다.

컴퓨팅 하드웨어

워크로드 요구사항에 따라 하드웨어 구성을 선택합니다. 컴퓨팅 최적화, 메모리 최적화 및 기밀 컴퓨팅 하드웨어의 가용성은 지역, 서비스 계층 및 컴퓨팅 계층에 따라 달라집니다.



2. 트러블슈팅 - LLM Training

필수:

인스턴스 시리즈: NC_T4_v3-series

인스턴스: NC4as T4 v3: 4 vCPUs, 28 GB RAM, 180 GB Temporary storage, US\$0.8...

Virtual Machines: 4

시간: 1

절약 옵션

가격 책정 모델을 탐색하여 Azure 비용을 최적화합니다.

자세한 정보

Compute (NC4as T4 v3)

OS (Windows)
 라이선스 포함
 Azure Hybrid Benefit

절약 플랜

1년 결감 플랜 (~32% 할인)

3년 결감 플랜 (~53% 할인)

예약

1년 예약 (~41% 할인)

3년 예약 (~62% 할인)

US\$2.59
월별 평균
(US\$0.00 신불로 요금이 청구됨)

US\$0.74
월별 평균
(US\$0.00 신불로 요금이 청구됨)

= US\$3.32
월별 평균
(US\$0.00 신불로 요금이 청구됨)

Azure MV + GPU

T4 * 4 예상 비용(학습 시간 40시간
예상) = 약 132 달러

G-NAHP-80
A100 80GB PCIe
 라이선스 포함

- CPU 16 vCore
- Memory 192 GiB

생성된 인스턴스 0/1

₩2,000/1시간마다

G-NAHPM-40
A100 80GB PCIe MIG 3g-40GB
 라이선스 포함

- CPU 8 vCore
- Memory 96 GiB

생성된 인스턴스 0/1

₩1,380/1시간마다

Elice A100

A100 비용(학습 예상
시간 2~3시간) = 6천원



배포 이슈 존재 MV
리소스와 같음

모델 학습을 위해 세팅했던 환경들- cost에 대한 것들
(고려했던 사항들)

ML 컴퓨팅 > 1장짜리 gpu > elice cloud > 4장 짜리 gpu

3. 트러블슈팅 - 디버깅 및 문제해결

A100 메모리에 풀 적재 되지 않는 문제

배치 설정 문제로 배치를 적게 잡아둬서
풀 적재 되지 않았음.

| scripts/bg_train.sh

- 기존

```
--per_device_train_batch_size=2 \
--gradient_accumulation_steps=64 \
```

Shell

- 변경

```
--per_device_train_batch_size=1 \
--gradient_accumulation_steps=128 \
```

Shell

```
Every 1.0s: nvidia-smi                                         9d9eb4564980: F
Fri Jun 27 14:41:35 2025
+-----+
| NVIDIA-SMI 535.183.06           Driver Version: 535.183.06   CUDA Version: 12.4 |
+-----+
| GPU  Name                  Persistence-M | Bus-Id     Disp.A  Volatile Uncorr. ECC | | | | | | |
| Fan  Temp      Perf          Pwr:Usage/Cap | Memory-Usage | GPU-Util Compute M. |
|          |          |          |          |          |          |          |          |
|-----+
| 0  NVIDIA A100 80GB PCIe    On        00000000:65:00.0 Off   0          |
| N/A  44C     P0          265W / 300W | 64256MiB / 81920MiB | 86%     Default |
|          |          |          |          |          |          |          |
+-----+
+-----+
| Processes:
| GPU  GI  CI      PID  Type  Process name          GPU Memory |
| ID   ID          ID   ID          Usage          |
|-----+
+-----+
```