

생성 AI 기반 TOPIK 문제 자동생성
및 모의학습 CBT 플랫폼 구현

Implementation of an AI-based TOPIK Question
Generation and Mock Test CBT Platform



신병근

김범수

허진영

지도교수 조준수

목 차

1. 서론.....	1
1.1. 연구 필요성	1
1.2. 기존 문제점	2
1.3. 연구 목표.....	2
2. 연구 배경.....	3
2.1. AI 기술의 발전.....	3
2.2. 기존 연구와의 차별성.....	3
3. 연구 내용	4
3.1. 문제 자동 생성 시스템.....	4
3.1.1. 데이터 크롤링	4
3.1.2. 프롬프트 엔지니어링.....	7
3.1.3. 대규모 언어 모델(LLM: Large Language Model)	8
3.1.4. 문제 생성 병렬 처리.....	9
3.1.5. RAG(Retrieval-Augmented Generation)	9
3.1.6. Model Serving.....	10
3.2. 모의시험 플랫폼 구현.....	10
3.2.1. 반응형 웹 기반 플랫폼	10
3.2.2. PBT, CBT 기반 모의고사	10
3.2.3. 생성 문제 검수.....	11
3.2.4. 회원 관리.....	11
3.2.5. 개발 언어 및 개발 도구.....	11

3.2.6. 시스템 구조	12
4. 연구 결과 분석 및 평가	25
4.1. 최종 완성 기능 목록	25
4.2. 성능 평가	26
4.2.1. 생성 품질	26
4.2.2. 생성 과정	28
4.2.1. 생성 비용	29
4.2.2. 생성 속도	30
4.3. 결과물	30
4.4. 테스트	35
5. 결론 및 향후 연구 방향	36
6. 참고 문헌	37

1. 서론

1.1. 연구 필요성

본 프로젝트는 TOPIK(Test of Proficiency in Korean) 시험을 준비하는 학습자들을 위해, 생성 AI 기술을 활용하여 문제를 자동 생성하고 모의 CBT(Internet-Based Test) 플랫폼을 구현하는 것을 목표로 하였다. 한국어에 대한 관심이 증가함에 따라 응시자 수가 늘어 기존의 종이 시험 및 수기 채점 형태에서 디지털화가 시급해졌기에 컴퓨터 기반의 TOPIK 문제 공부 및 모의 시험 플랫폼이 한국어 교육 현장에서 큰 필요성을 갖고 있다.

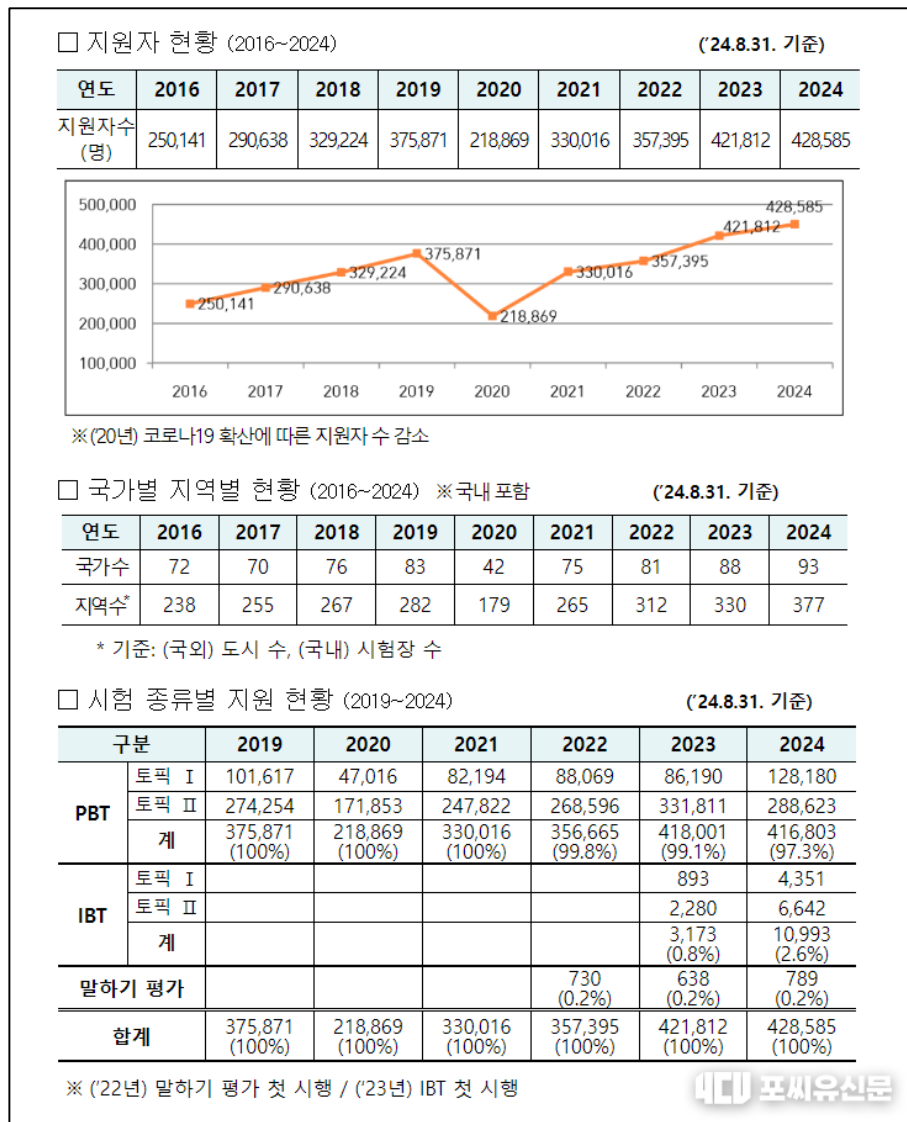


그림1. 한국어능력시험(TOPIK) 주요 통계 자료¹

¹ [한국어능력시험\(TOPIK\), 2025년부터 인터넷기반시험\(IBT\) 응시 기회 대폭 확대 \(4cu.com\)](https://www.4cu.com)

1.2. 기존 문제점

현재 한국어 능력시험을 준비하는 학습자들은 종이 형태의 교재와 기출문제 위주로 학습을 진행하지만, 2023년부터 **TOPIK 시험의 IBT 방식 전환²**으로 디지털 교육 자료와 교육 받을 장소가 부족하다는 점이 문제로 대두된다. 특히, 해외 수험생들은 시험을 보기 위해 이동 및 숙소 문제로 경제적 부담이 크고, 교육 자료에 대한 접근성 또한 떨어진다.

2023년 대륙별 지원 현황							
구분	아시아	북아메리카	남아메리카	아프리카	유럽	오세 아니아	전체
지원자수	406,182	1,630	2,160	327	10,186	591	421,812
비율	96.47%	0.39%	0.51%	0.08%	2.41%	0.14%	100%
* 출처 : 교육부, 국립국제교육원							

표1. 2023년 대륙별 TOPIK 지원 현황³

위 그림1과 표1에 따르면, TOPIC 응시자는 2020년 21만 8869명에서 2023년 42만 1812명으로 3년 동안 2배 이상으로 늘었다. 대륙별 지원 현황으로는 아시아 지역이 다수를 차지했다. 한국에서의 응시자를 제외한 24만 2648명 가운데 93.6%인 22만 7754명이 아시아에서 응시했다. 전 세계 한류동호회 ⁴ 가입인구 2억 2500만명 중 1억 4700만명 (65.3%)이 아시아 대륙에 있었다. 이렇듯, **아시아에서 한국어 학습에 대한 관심이 컸지만, 교육 시설은 이에 미치지 않는다**. 23년 교육통계연보에 ⁵ 따르면, 전 세계에 있는 한국어 교육시설 1433개 중 아시아에 위치한 교육 시설은 20.2%인 290개에 불과했다.

1.3. 연구 목표

따라서 본 과제는 TOPIK 시험 준비를 효율적으로 지원하기 위한 플랫폼을 개발하는 것을 목표로 하였다. 구체적으로는 **AI 기반 문제 자동 생성**과 **모의시험 응시**를 지원하여, 학습자의 편의를 증대시키고 학습의 질을 높임과 동시에, 교육 환경과 교육 자료의 확대와 더불어 한국어 학습자 데이터를 확보함에 의의가 있다.

² [한국어능력시험\(TOPIK\) 디지털 전환 계획](#)

³ [한국어시험 응시자 10명 중 9명은 아시아, 교육시설은 절반 이상 북미에 | 서울신문 \(seoul.co.kr\)](#)

⁴ [KF통계센터-전세계 한류동호회원 현황 통계](#)

⁵ [교육통계연보 | 교육통계서비스 \(kedi.re.kr\)](#)

2. 연구 배경

2.1. AI 기술의 발전

최근 대규모 언어 모델(LLM: Large Language Model)의 발전으로, 자동 문제 생성 기술이 에듀 테크 분야에서 큰 가능성을 보이고 있다. 이에 GPT-3와 같은 모델로 다양한 유형의 문제를 인간과 비슷한 수준으로 생성하거나 변형할 수 있는 능력을 증명한 선행 연구들이 존재한다. [1, 2]

2.2. 기존 연구와의 차별성

기존 연구에서는 AI 기반 문제 생성이 주로 특정 도메인, 예를 들어 의학 교육[3]이나 수학 문제 생성[4]의 제한된 분야에 집중되었다. 그러나 이러한 연구들은 대부분 단일 유형의 문제 생성에 국한되었고, 언어 교육과 같은 여러 주제의 지문에서 다양한 유형의 문제를 다루는 데에는 한계가 있었다[5].

본 프로젝트에서는 한국어 능력시험(TOPIK)에 특화된 문제 생성 모델을 개발하였으며, 특히 문법, 빈칸 채우기, 주제 찾기와 같은 다양한 유형의 문제를 자동으로 생성할 수 있는 기능을 구현하여 기존 연구와 차별화된다.

뿐만 아니라, RAG(Retrieval-Augmented Generation)와 LLM의 프롬프트 기반 Few-Shot Learning을 응용하여 기존 지문의 스타일을 유지하면서도 새로운 주제의 지문과 선택지를 구성하는 기술을 개발하여 지문 구조를 원본 시험과 다르지 않게 크게 바꾸지 않으면서도 다양성을 확보한 것이 특징이다. 또한 쓰기 영역에서 LLM 프롬프팅을 활용한 자동 채점 기능을 구현하여, 기존의 사람이 채점하는 주관적인 채점 방식에서 채점 기준에 따라 일관성을 유지하는 채점 시스템을 구현하였다.

최종적으로는, 문제은행 시스템을 통해 생성된 문제를 축적하고 재사용할 수 있도록 하여, 학습자가 언제든지 다양한 유형의 문제를 접할 수 있는 환경을 제공함으로써 기존 연구보다 더 확장성과 실용성을 갖춘 모델을 제시한다.

3. 연구 내용

3.1. 문제 자동 생성 시스템

AI 기반 문제 생성 시스템은 기존 문제 스타일을 벗어나지 않기 위해 기출문제를 기반으로 새로운 문제를 생성하는 방식으로 설계되었다.

가장 먼저 생성하게 될 주제를 요청 받거나, 따로 없다면 국립국어원에서 자주 사용되는 한국어 5800자 데이터셋에서⁶ 단어를 선정하고 명사가 아닌 형용사나 동사가 선택될 경우 LLM을 통해 그에 맞는 적절한 주제가 생성된다. 이 경우, 교육적 도메인 특성상 위험한 주제를 회피하기 위한 프롬프트 튜닝을 하였다.

그 다음 RAG(검색 증강 생성) 기술로 생성하기 전 생성 요청된 비슷한 유형의 문제를 탐색하여 비슷한 구조의 지문을 가져오게 된다. 이렇게 비슷한 지문 2-3가지를 추출한 뒤 LLM의 프롬프트 기반 Few-Shot Learning을 활용하여 선택된 주제와 결합해 비슷한 구조의 새로운 주제로 구성된 지문이 생성된다.

이후 요청된 문제나 참고한 문제 유형에 맞게 선택지를 구성하고 고추론적 사고가 가능한 LLM(GPT-4등)을 용해 지문의 구성 및 오답과 정답을 검증한다.

이미지가 포함된 문제의 경우, GPT4o의 멀티모달 기능을 이용하기엔 이미지 Input 토큰이 상당히 많이 소요되고 내용의 정확도가 중요해 Google Cloud Vision API를 선택해 OCR를 수행하고 직접 텍스트를 읽어와 문제와 선택지를 구성하게 된다.

이러한 방식을 통해 각 시험(토픽I, 토픽II)에 맞게 문제 유형별로 새로운 문제를 지정된 개수만큼 생성하고, 문제은행에 등록하여 이를 기반으로 한 또 다른 문제 생성이 가능하도록 하였다.

3.1.1. 데이터 크롤링

생성을 위한 기반 문제 데이터베이스를 구축하기 위해서 한국어능력시험(TOPIK) 공식 홈페이지에서 공개한 기출문제⁷를 크롤링하였다. 최근 10개년(14~24년) 토픽 I,II 듣기, 읽기 시험(총 40세트)를 확보하여 초기 문제 데이터베이스를 구축하였다.

문제 데이터 구축 시 유형과 구조를 분석하여 정리한 내용은 다음의 표와 같다.

⁶ 부록:자주 쓰이는 한국어 낱말 5800 - 위키낱말사전 (wiktionary.org)

⁷ [한국어능력시험\(TOPIK\) 기출문제](#)

		유형	데이터 명	토픽I
토픽I	1	물음에 맞는 대답 고르기	listen_1_problem_type_1	1-4번
듣기	2	이어질 말	listen_1_problem_type_2	5-6번
30문항	3	알맞은 것 고르기	listen_1_problem_type_3	7-14번
	4	알맞은 그림 고르기	listen_1_problem_type_4	15-16번
	5	같은 내용 고르기	listen_1_problem_type_5	17-24번
	6	무엇을 하는지+같은 내용 고르기	listen_1_problem_type_6	25-28번
	7	이유+같은 내용 고르기	listen_1_problem_type_7	29-30번
토픽I	8	알맞은 것 고르기	reading_1_problem_type_1	31-39번
읽기	9	맞지 않은 것 고르기(그림/표)	reading_1_problem_type_2	40-42번
40문항	10	내용이 같은 것 고르기	reading_1_problem_type_3	43-45번
	11	중심 생각 고르기	reading_1_problem_type_4	46-48번
	12	빈칸에 들어갈 말 +같은 것 고르기	reading_1_problem_type_5	49-56번
	13	순서 나열	reading_1_problem_type_6	57-58번
	14	빈칸에 들어갈 말+같은 것 고르기	reading_1_problem_type_7	63-64번
	15	글쓴이의 의도 파악+같은 것 고르기	reading_1_problem_type_8	59-62,65-70번
토픽II	16	알맞은 그림 고르기	listen_2_problem_type_1	1-3번
듣기	17	이어질 수 있는 말	listen_2_problem_type_2	4-8번
50문항	18	이어서 할 행동	listen_2_problem_type_3	9-12번
	19	내용 일치	listen_2_problem_type_4	13-16번
	20	중심 생각	listen_2_problem_type_5	17-20번
	21	중심 생각+맞는 것 고르기	listen_2_problem_type_6	21-22,25-26,33-34번
	22	무엇을 하는지+맞는 것 고르기	listen_2_problem_type_7	23-24,35-36번

	23	말하는 의도 파악+맞는 것 고르기	listen_2_problem_type_8	27-28번,
	24	직업 고르기+맞는 것 고르기	listen_2_problem_type_9	29-30번
	25	생각 고르기+태도	listen_2_problem_type_10	31-32번
	26	공식적 대화 (교양)	listen_2_problem_type_11	37-38번
	27	공식적 대화 (대담)	listen_2_problem_type_12	39-40,47-48번
	28	공식적 대화 (강연)	listen_2_problem_type_13	41-42,45-46,49-50번
	29	공식적 대화 (다큐멘터리)	listen_2_problem_type_14	43-44번
토픽II	30	알맞은 문법 고르기	reading_2_problem_type_1	1-2번
읽기	31	알맞은 유의 문법,표현 고르기	reading_2_problem_type_2	3-4번
50문항	32	광고,소재	reading_2_problem_type_3	5-8번
	33	같은 것 고르기(도표,그래프)	reading_2_problem_type_4	9-12번
	34	순서 배열	reading_2_problem_type_5	13-15번
	35	빈칸 채우기	reading_2_problem_type_6	16-22번
	36	빈칸+같은 것/중심 생각	reading_2_problem_type_7	23-26번
	37	심정+같은 것 고르기	reading_2_problem_type_8	27-28번
	38	신문 기사 제목 고르기	reading_2_problem_type_9	29-30번
	39	같은 것 고르기	reading_2_problem_type_10	32-34번
	40	주제 찾기	reading_2_problem_type_11	35-38번
	41	문맥에 맞는 위치 고르기	reading_2_problem_type_12	39-41번
	42	소설	reading_2_problem_type_13	42-43번
	43	설명문/논설문	reading_2_problem_type_14	44-45번
	44	정보(순서 배열)	reading_2_problem_type_15	46-47번
	45	종합(논설문)	reading_2_problem_type_16	48-50번

표2. TOPIK 시험별 유형 분석표

listen/reading_1/2_problem_type_N

※ [1 ~ 4] 다음을 듣고 <보기>와 같이 물음에 맞는 대답을 고르십시오.

problem:string

<보기>

가: 물을 마셔요?
나: _____

① 네, 물을 마셔요. ② 네, 물이 적어요.
③ 아니요, 물이에요. ④ 아니요, 물이 좋아요.

example:json

example은 유형에 따라 있을수도 없을 수도 있음, 형태는 유형마다 다름.

1. (4점) questionNumber:int, question:string, score:int

① 네, 학생이에요. ② 네, 학생이 없어요.
③ 아니요, 학생이 와요. ④ 아니요, 학생이 많아요.

answer:string[]

2. (4점)

① 네, 책을 싫어해요. ② 네, 책이 아니에요.
③ 아니요, 책이 있어요. ④ 아니요, 책을 안 읽어요.

answer:string[]

question에도 보기가 유형에 따라 있을수도 없을수도 있음

questions:
question[]

그림2. TOPIK 문항 형태 자료 구조

3.1.2. 프롬프트 엔지니어링

최종적으로 생성할 문제는 읽기 유형 시험을 타게팅 하여, 읽기 I 문제 유형 10개와 읽기 II 문제 유형 17개, 쓰기 문제 유형 3개, 쓰기 채점에 대한 프롬프트 엔지니어링을 진행하였다. 이 과정에서 활용되는 기법은 다음과 같다.

LLM(Large Language Model, 대규모 언어 모델)에서의 Few-shot prompting는 GPT-3나 GPT-4 같은 모델에 소수의 예시(보통 하나에서 다섯 사이)를 제공하여 그 반응을 안내하거나 특정 작업을 가르치는 기술을 말한다.

Few-shot prompting에서는 모델에게 주어진 예시들이 패턴을 보여주거나 특정한 반응 방식을 가르치기 위해 선택된다. 그런 다음 모델은 이 예시들을 참조로 사용하여 자신의 반응에서 기대되는 것을 이해한다. 이는 단순한 지시만으로는 모델이 의도된 작업을 충분히 이해하기 어려운 복잡하거나 미묘한 작업에 특히 유용하다.

이러한 방식으로 한국어 문제 생성 중, 기존 시험과 크게 다르지 않게 지문의 스타일을 유지하는 역할을 한다. 이 작업을 수행하지 않고 단순한 프롬프팅으로 문제를 생성할 경우, 지문의 퀄리티가 일정하지 않고 문제로서 부적합한 지문이 생성될 가능성이 높아진다.

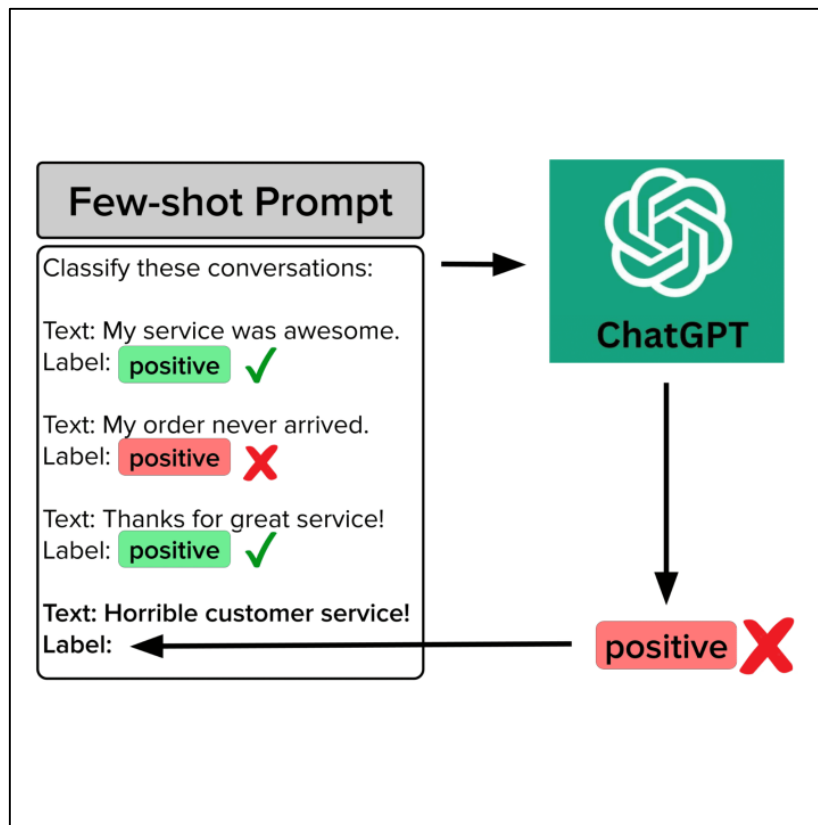


그림3. Language Models are Few-Shot Learners [1]

3.1.3. 대규모 언어 모델(LLM: Large Language Model)

최종적으로 **성능, 응답시간, 비용**을 종합적으로 고려하여 생성을 위해 선택한 모델은 OpenAI의 최신 모델 **gpt-4o-mini**이다. MMLU은 LLM의 추론 성능을 평가하는 대표적인 벤치마크(척도) 중 하나로, 언어 시험 문제 생성 능력을 평가하는 데 가장 적합하다. 문제 생성은 입출력 토큰이 많이 발생하기에 추론 성능이 살짝 낮더라도 비용이 가장 낮은 GPT-4o mini에 프롬프트 튜닝, COT(Chain-of-Thought), Few-Shot Learning을 통해 보강을 적용하여 사용하였고, 문제가 올바르게 생성 되었는지 검증하는 모델은 추론 성능이 높은 GPT-4o, GPT-4를 Ensemble하여 교차 검증 하는 형태로 사용하였다.

Model	Performance(MMLU)	Throughput(/sec) ⁸	Cost(/1M tokens)
GPT-4o	88.7	123 tokens	\$12.5
GPT-4o mini	82.0	100 tokens	\$0.75
GPT-4 Turbo	86.5	36 tokens	\$40.0

표3. OPEN AI LLM 비교

⁸ [Comparison of AI Models across Quality, Performance, Price | Artificial Analysis](#)

3.1.4. 문제 생성 병렬 처리

비동기 처리가 가능한 파이썬 서버를 통해서 문제 생성 병렬 처리를 구현하였다. 이를 통해 문제 생성 시간을 단축하였다. 10개 유형이 있고, 하위에 2개씩의 문제가 있다고 가정했을때, LLM의 생성 속도 한계상 각 문제당 생성 시간이 평균적으로 10초 정도일 경우 일반적인 동기 방식으로 작동하면 $10 \times 2 \times 10 = 200$ 초라는 시간이 걸린다.

본 과제에서는 이를 객체 지향적 구조를 도입하여 모든 문제 생성 유형을 독립시키고, 파이썬의 asyncio 모듈로 각 유형의 생성 Task를 비동기로 처리하며 각 문제의 생성에는 RAG, 자연어처리 등의 처리가 필요한 CPU-Bound작업이 수행되기에 파이썬의 GIL(Global Interpreter Lock)을 회피하고자 multiprocessing 모듈을 적용하여 동시에 여러 문제를 생성한다. 이러한 경우, 최종적으로 가장 늦게 생성된 유형의 생성 시간이 최대 작업 시간이 되며 위와 같은 경우 200초에서 약10~20초 정도의 시간으로 단축된다.

3.1.5. RAG(Retrieval-Augmented Generation)

LLM 관련 프레임워크인 LangChain, NoSQL기반의 MongoDB를 사용하여 검색 증강 시스템을 구현한다. 수집한 문제를 별도로 백엔드 RDBMS에 정량화하여 저장한 뒤, OPENAI text-embedding-3-large 모델을 이용해 벡터 임베딩을 적용해둔 것을 MongoDB로 보관한다. 이후 쿼리에 따라 적절한 지문을 유사도 순으로 추출한다.

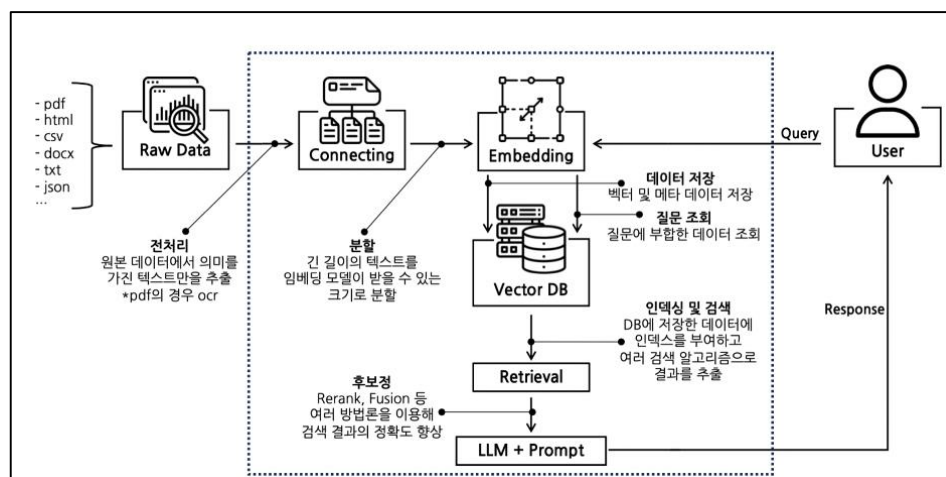


그림4. RAG(Retrieval-Augmented Generation) 구조

3.1.6. Model Serving

ASGI Worker를 통해 Python Fast API의 병렬 사용과 더불어 비동기 통신이 가능한 생성 AI용 API 서버를 구현하여 문제 생성기를 Serving하고 모의시험 웹 플랫폼 백엔드 서버와 연동한다. 파이썬에서 대표적으로 사용되는 웹 서버 프레임워크로 Django, Flask가 존재하는데 Fast API를 선택한 이유는, AI 모델 서빙 시에는 요청 처리 속도가 중요하며, 특히 대량의 요청을 병렬 처리해야 할 때 성능 차이가 두드러진다. 이때, Fast API는 비동기 코루틴을 쉽게 활용할 수 있어 병목 현상을 줄여준다. Django 또한 이를 지원하지만 고성능과 비동기 처리가 중요한 경우 Django는 오버헤드가 크다. 특히, AI 모델 같은 속도와 자원에 민감한 작업에는 불필요한 기능들이 함께 동작하여 성능 저하가 발생할 수 있다. Flask는 가벼운 마이크로프레임워크로, 불필요한 기능 없이 필요한 만큼만 기능을 추가할 수 있어 매우 유연하지만, 비동기 처리를 기본적으로 지원하지 않으며 너무 자유롭기에 기본적으로 지원하는 기능이 적다.

따라서 본 과제에서 개발의 편의와 AI모델 서빙으로 Fast API 배포가 최적으로 보인다.

3.2. 모의시험 플랫폼 구현

3.2.1. 반응형 웹 기반 플랫폼

모의시험 플랫폼은 반응형 웹 기반으로 구현되어 컴퓨터로 접근이 힘든 지역의 학습자들 또한 사용할 수 있게 인터넷을 사용할 수 있는 모든 디바이스(데스크톱, 태블릿, 휴대폰)에서 직접 응시할 수 있도록 설계되었다. 사용자는 원하는 문제 유형을 선택하여 시험을 구성할 수 있으며, 응시 후 자동 채점 및 통계 분석 기능을 통해 자신의 학습 성취도를 파악할 수 있다.

3.2.2. PBT, CBT 기반 모의고사

모의고사는 인터넷 응시 뿐만 아니라 각지의 학원에서 혹은 자체적으로 PBT 형태의 종이 시험을 연습할 수 있도록 프린트가 가능한 PDF 생성 기능을 지원한다. 해당 시험지는 시험지 커버, 정답지 여부, 1단 2단 레이아웃 배치를 지원하여 학습자의 편의를 높였다.

CBT는 웹UI에서 각 유형에 맞게 공식 토픽 시험과 유사한 UI를 지원하며, OMR카드와 남은 시간 타이머가 시험지 옆에 함께 배치되어 응시자가 선택한 정답 및 진도율을 파악할 수 있게 하였다. 또한 생성된 문제 은행의 문제들을 기반으로 셔플 기능을 지원하여 사용자가 늘 새로운 문제를 접하면서 학습할 수 있게 하였고, 사용자가 직접 학습하고 싶은 유형을 선택하여 시험지를 구성할 수 있도록 구현한다. 듣기 시험은 듣기 파일 재생

을 지원하며 시험의 공정성을 위해 원하는 위치 이동을 없애고 2배속을 구현해 두었다.

또한 유저가 등록할 수 있는 특별 커스텀 시험을 지원하여 자유도를 높였다.

채점시 각 문제와 OMR카드에 O,X 표시로 정답 유무를 확인할 수 있으며 OMR카드의 번호를 눌러 빠른 스크롤 이동을 지원한다.

3.2.3. 생성 문제 검수

생성형 AI의 경우 아무리 정확도가 높은 최신 모델일지라도 할루시네이션으로 인해 결과물에 대한 인간의 검증이 반드시 필요하다. 특히 교육, 문제 도메인 특성상 더욱 중요하다. 따라서 생성된 문제 은행에 접근하여, 전문가나 관리자가 쉽게 문항을 수정할 수 있는 검수 페이지를 구현하여 도입한다.

3.2.4. 회원 관리

본 과제의 플랫폼은 한국어 학습자 데이터 확보에도 목적이 있으므로 인증된 사람들이 이용 가능하도록, 회원을 도입하지만 복잡한 회원 가입 절차를 생략하고 글로벌하게 이용될 수 있도록 구글 로그인 기반 회원 가입 및 로그인을 구현한다. 이를 통해 국적, 생년월일, 이름 등의 기본적인 정보를 취득하고 추후 데이터 분석에 활용될 수 있도록 한다. 또한 학습 관리에도 중점을 두어 학습자들이 시험을 시도한 기록이나, 성적 결과, 해당 학습자가 풀었던 유형 및 통계를 저장하여 사용자의 학습 데이터를 기록한다.

3.2.5. 개발 언어 및 개발 도구

3.2.5.1. 개발 언어

FrontEnd	Html, Css, JavaScript
BackEnd	Java
AI Server	Python

3.2.5.2. 개발 도구

FrontEnd	Vite, React, Tailwind, with ESLint
BackEnd	Spring, PostgreSQL, Nginx
Infra	Docker, Redis, AWS (VPC, EC2, S3, RDS)
AI Server	ASGI worker, FastAPI, Langchain, Mongo, Google Cloud Platform AI
LLM	GPT-4-turbo, GPT-4o, GPT-4o-mini

3.2.6. 시스템 구조

3.2.6.1. 프론트엔드(Front-end)

Vite + React JS + tailwind를 사용해 사용자 친화적인 웹 기반 CBT(Computer-Based Test) 시스템을 구현한다.

Vite는 전통적인 번들러보다 훨씬 빠른 개발 환경을 제공하는데, 빠른 HMR(Hot Module Replacement) 덕분에 시행착오와 피드백이 많은 초기 애플리케이션 개발에 적합하다. React는 컴포넌트 기반으로 UI를 설계할 수 있어 복잡한 화면을 재사용 가능한 컴포넌트로 분리하는 덕분에 각 화면을 모듈화 하고 유지보수와 확장성이 뛰어나다.

Tailwind는 유틸리티 클래스 기반의 CSS 프레임워크로, 빠르게 UI를 구성하고 디자인 일관성을 유지할 수 있어 개발 속도와 편의가 증가하고, 반응형 웹의 구성이 굉장히 쉬워지는 장점이 있다.

본 과제의 웹은 크게 시험지, 문제 유형별 컴포넌트, 메인 랜딩페이지 구조로 구성되며, 로그인 및 회원 관리를 추가하여 프로젝트의 완성도를 높인다.

3.2.6.2. 백엔드(Back-end)

본 프로젝트의 인프라 시스템은 **AWS 클라우드 서비스**를 기반으로 설계되어, **확장성**과 **유연성**을 극대화하면서 **비용 효율성**을 확보하였다. 주요 구성 요소는 다음과 같다:

1. 서버 측 애플리케이션

- **AWS EC2 인스턴스**를 활용하여 여러 개의 **Docker 컨테이너**로 애플리케이션을 배포한다. 이를 통해 각각의 애플리케이션 모듈이 독립적으로 운영될 수 있으며, 서버 간 확장성이 용이하다.
- 백엔드 서버는 **Spring 프레임워크**로 구성되어 있으며, **Restful API 통신**을 이용한다. 또한, 멀티스레딩을 활용하여 **트래픽 처리 성능**을 최적화하였다.
- Spring은 IoC (Inversion of Control) 컨테이너와 AOP (Aspect-Oriented Programming) 기능을 제공하여 테스트 코드 작성과 유지보수가 용이하다.
- 각 레이어(서비스, 컨트롤러)에 회원 권한에 따른 접근 인가를 구현하였으며 그 과정에서 중복된 회원 정보 조회를 피하기 위해 ThreadLocal을 이용하여 첫 번째 조회시 회원 정보를 ThreadLocal에 저장하여 그 이후에 필요한 회원 데이터를 ThreadLocal에서 가져오게 하였다.

- 회원을 도입하게 되면서 각 회원의 학습의 통계 처리를 하는것이 요구되어 현재 통계 데이터 조회를 위해 시험 기록 데이터를 조회후 전체 데이터에서 통계를 내는 방식이 비효율적이고 많은 트래픽이 올 시 병목현상이 생길 우려가 있어 비정규화와 통계 테이블을 생성하여 시험을 쳤을 때마다 그 기록을 일일이 기록을 해두도록 설계를 하였다.
- Mockito를 활용하여 서비스 layer와 데이터베이스 layer에 대한 단위 테스트를 진행하였으며, 이 과정에서 Mocking과 Stubbing을 적절히 사용하여 각 레이어의 독립적인 역할을 테스트하였다. 이를 통해 외부 의존성을 최소화하고, 테스트 범위 내에서 원하는 동작이 잘 작동하는지를 확인함으로써 코드의 신뢰성과 유지보수성을 향상시켰다

Method	URL	Description	Parameters
PUT	/group/{groupId}	그룹 업데이트 - 이름	groupId
PUT	/center/{centerId}	센터 정보 업데이트 - 이름, 주소, 국가	centerId
DELETE	/center/{centerId}	센터 삭제	centerId
POST	/problem	문제 등록	None
POST	/problem/random	문제타입에 맞는 랜덤 문제 조회	None
POST	/member/{memberId}/detail	회원 상세 정보 추가 - 국가, 성별, 생일	memberId
POST	/member/group/offer/{offerId}	그룹 가입 승인	offerId
POST	/member/google	구글 로그인	None
POST	/member/center/offer/{offerId}	센터 가입 승인	offerId
POST	/group/offer	그룹 가입 오퍼 등록	None
GET	/group/	자신의 센터 그룹 목록 조회	None

POST	/group/	그룹 생성	None
GET	/exam	시험 전체 조회	None
POST	/exam	시험 등록	None
GET	/exam/generated	generated 시험 전체 조회	None
POST	/exam/generated	Personal 시험 생성	None
POST	/exam/department/record	소속별 기간 기록 조회	None
GET	/exam-answer/{examId}	examId 시험 정답 조회	examId
POST	/exam-answer/{examId}	examId 시험 정답지 제출 및 결과 저장	examId
POST	/credit/	크레딧을 부여한다	None
GET	/center/offer	자신의 센터 가입 오퍼 목록 조회	None
POST	/center/offer	센터 가입 오퍼 등록	None
GET	/center/	센터 목록 조회	None
POST	/center/	center 생성	None
PATCH	/member/auth	회원 권한 변경	None
GET	/exam/{examId}	ExamId 를 이용한 시험 조회	examId
DELETE	/exam/{examId}	examId 시험 삭제	examId
PATCH	/exam/{examId}	examId 시험 업데이트 - modify	examId
GET	/problem/{problemId}	문제 조회	problemId
GET	/member/{memberId}	회원 정보 조회	memberId
GET	/member/validate	토큰 유효 여부 조회	None

GET	/exam/{examId}/group/{groupId}	examId 시험 그룹별 결과 조회	examId, groupId
GET	/exam/{examId}/center/{centerId}	examId 시험 센터별 결과 조회	examId, centerId
GET	/exam/topik2-reading	topik2-reading 시험 전체 조회	None
GET	/exam/topik2-listening	topik2-listening 시험 전체 조회	None
GET	/exam/topik1-reading	topik1-reading 시험 전체 조회	None
GET	/exam/topik1-listening	topik1-listening 시험 전체 조회	None
GET	/exam-record/{memberId}	memberId 시험 결과 조회	memberId
GET	/credit/used	사용한 크레딧 조회	None
GET	/credit/unused	사용하지 않는 크레딧 조회	None
GET	/analyze/{memberId}	member 학습 분석 조회	memberId
DELETE	/group/{centerId}	그룹 삭제	centerId
DELETE	/center/offer/{offerId}	센터 가입 오퍼 삭제	offerId

표4. API 명세

2. 데이터베이스

- 기출문제 데이터를 크롤링하여 **Amazon RDS의 PostgreSQL**을 사용해 **Master-Slave** 구조로 저장하였으며, 이를 통해 고가용성을 유지한다.
- PostgreSQL은 오픈소스 데이터베이스이어서 라이선스 비용이 발생하지 않고 ACID 트랜잭션을 지원하여 데이터의 일관성과 무결성을 보장한다.
- 정렬이 필요한 테이블의 경우에는 인덱스를 생성하여 데이터를 더욱 효율적으로 검색할 수 있도록 하였다.

- 45개의 서로 다른 문제 유형들을 추상화를 통해 5개의 테이블로 만들었으며 이를 통해 데이터 구조를 단순화하여 유지보수에 편리하게 하였다.
- 현재 데이터베이스에서 데이터 삭제시 바로 삭제되기 때문에 데이터 복구를 위해 각 테이블에 is_deleted 칼럼을 추가하여 데이터 논리적 삭제를 도입하였으며 추후 batch와 scheduler 를 도입하여 일정기간 이후에 데이터가 완전 삭제가 되도록 구현할 예정이다.

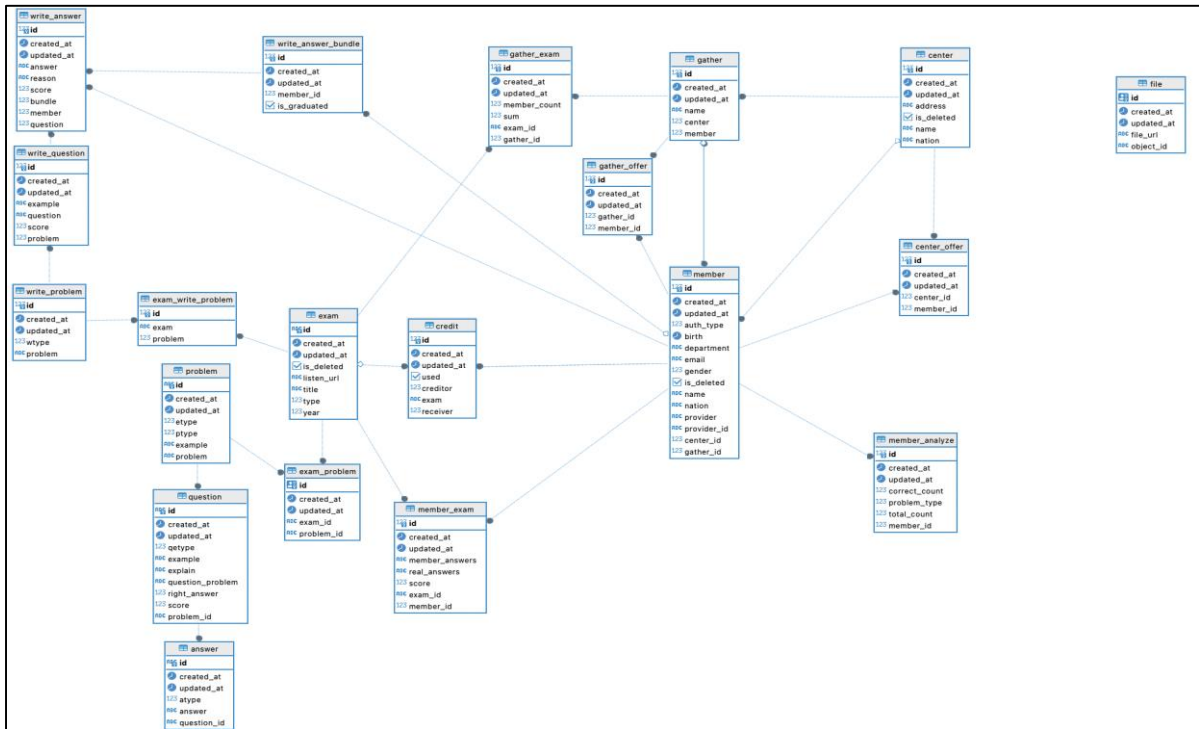


그림5. 데이터베이스 ERD

테이블 명	필드	타입	제약	설명
CENTER	id	bigint	NOT NULL	센터 데이터 구분을 위한 id
	address	varchar(255)		센터의 주소
	name	varchar(255)		센터의 이름
	nation	varchar(255)		센터의 국가
	created_at	timestamp p(6) with time zone		센터 생성일
	updated_at	timestamp p(6) with time zone		센터 정보 업데이트일

	is_deleted	boolean		센터 삭제여부
CENTER_OFFER	id	bigint	NOT NULL	센터 가입신청 데이터 구분을 위한 id
	center_id	bigint	NOT NULL	가입신청한 센터의 id
	member_id	bigint	NOT NULL	가입신청한 멤버의 id
	created_at	timestamp p(6) with time zone		가입신청 생성일시
	updated_at	timestamp p(6) with time zone		가입신청 업데이트 일시
CREDIT	id	bigint	NOT NULL	크레딧 데이터 구분을 위한 id
	creditor	bigint	NOT NULL	크레딧을 부여한 멤버의 id
	receiver	bigint	NOT NULL	크레딧을 받은 멤버의 id
	used	boolean default false	NOT NULL	크레딧 사용 여부
	exam	bigint		크레딧을 사용한 exam의 id
	created_at	timestamp p(6) with time zone		크레딧 생성 일시
	updated_at	timestamp p(6) with time zone		크레딧 업데이트 일시
EXAM_PROBLEM	id	uuid	PRIMARY KEY NOT NULL	exam과 problem의 관계 데이터 구분을 위한 id
	exam_id	varchar(255)	NOT NULL	exam의 id
	problem_id	varchar(255)	NOT NULL	problem의 id
	created_at	timestamp p(6) with time zone		관계 생성일시
	updated_at	timestamp p(6) with time zone		관계 업데이트 일시
GATHER_EXAM	id	bigint	PRIMARY	모임의 시험 데이터

	gather_id	bigint	RY KEY NOT NULL NOT NULL	터 구분을 위한 id 모임의 id
	exam_id	varchar(255)	NOT NULL	시험의 id
	created_at	timestamp p(6) with time zone		모임의 시험 생성 일시
	updated_at	timestamp p(6) with time zone		모임의 시험 업데이트 일시
EXAM_WRITE_PROBLEM	id	bigint	PRIMARY KEY NOT NULL	exam과 write_problem의 관계 데이터 구분을 위한 id
	exam	varchar(255)	NOT NULL	exam의 id
	problem	bigint	NOT NULL	write_problem의 id
	created_at	timestamp p(6) with time zone		관계 생성일시
	updated_at	timestamp p(6) with time zone		관계 업데이트 일시
EXAM	id	varchar(255)	PRIMARY KEY NOT NULL	exam 데이터 구분을 위한 id
	title	varchar(255)	NOT NULL	시험의 이름
	listen_url	varchar(255)		시험의 듣기파일
	type	smallint		시험의 타입 (topik1,topik2,writing)
	year	integer		시험의 생성년도
	created_at	timestamp p(6) with time zone		데이터의 생성일시
	updated_at	timestamp p(6) with		데이터의 업데이트 일시

	time zone			
	is_deleted	boolean		시험의 삭제여부
MEMBER_EXAM	id	bigint	PRIMARY KEY NOT NULL	멤버의 시험 기록 데이터를 위한 id
	member_answers	varchar(255)		멤버의 선택한 정답 목록
	real_answers	varchar(255)		시험의 실제 정답 목록
	score	integer		시험 점수
	exam_id	varchar(255)	NOT NULL	시험의 id
	member_id	bigint	NOT NULL	멤버의 id
	created_at	timestamp p(6) with time zone		기록 생성일시
	updated_at	timestamp p(6) with time zone		기록 업데이트 일시
GATHER	id	bigint	PRIMARY KEY NOT NULL	모임 데이터를 구분할 위한 id
	name	varchar(255)		모임의 이름
	center	bigint	NOT NULL	센터의 이름
	member	bigint		모임을 생성한 사람
	created_at	timestamp p(6) with time		모임 생성일시
	updated_at	timestamp p(6) with time		모임 업데이트 일시
GATHER_OFFER	id	bigint	PRIMARY KEY NOT NULL	모임 가입신청 데이터를 구분할 위한 id
	gather_id	bigint	NOT NULL	모임의 id

	member_id	bigint	NOT NULL	모임 신청한 멤버 의 id
	created_at	timestamp p(6) with time		모임 신청의 생성 일시
	updated_at	timestamp p(6) with time		모임 신청 업데이트 일시
MEMBER_ANALYZE	id	bigint	PRIMA RY KEY NOT NULL	멤버별 문제 유형 분석 데이터를 구분 을 위한 id
	correct_count	integer		문제 유형의 맞은 갯수
	problem_type	smallint		문제 유형
	total_count	integer		문제 유형의 전체 갯수
	member_id	bigint	NOT NULL	멤버의 id
	created_at	timestamp p(6) with time		분석의 생성일시
	updated_at	timestamp p(6) with time		분석의 업데이트 일시
MEMBER	id	bigint	PRIMA RY KEY NOT NULL	멤버 데이터 구분 을 위한 id
	auth_type	smallint	NOT NULL	멤버의 권한 (student,teacher,c enter등등)
	email	varchar(2 55)	NOT NULL	멤버의 이메일
	is_deleted	boolean	NOT NULL	멤버의 삭제 여부
	name	varchar(2 55)	NOT NULL	멤버의 이름
	nation	varchar(2 55)		멤버의 국가
	provider	varchar(2 55)		멤버가 사용한 로 그인 제공자(구글)
	provider_id	varchar(2 55)		로그인 제공자가 부여한 id
	birth	date		생일

	gender	smallint		성별
	department	varchar(255)		임시 column 소속 센터
	center_id	bigint		소속 센터의 id
	gather_id	bigint		소속 모임의 id
	created_at	timestamp p(6) with time		멤버 생성일시
	updated_at	timestamp p(6) with time		멤버 업데이트 일시
ANSWER	id	varchar(255)	PRIMARY KEY NOT NULL	정답 데이터 구분을 위한 id
	atype	smallint		정답 데이터의 타입
	answer	varchar(255)		정답 데이터
	question_id	varchar(255)	NOT NULL	소속된 question의 id
	created_at	timestamp p(6) with time		정답 생성 일시
	updated_at	timestamp p(6) with time		업데이트 일시
PROBLEM	id	varchar(255)	PRIMARY KEY NOT NULL	problem 데이터 구분의 위한 id
	etype	smallint		예시의 데이터 타입
	ptype	smallint		problem의 타입 (45가지 문제타입)
	example	text		예시 데이터
	problem	varchar(255)		문제의 지문
	created_at	timestamp p(6) with time		문제 생성일시
	updated_at	timestamp p(6) with		문제 업데이트 일시

	time			
QUESTION	id	varchar(255)	PRIMARY KEY NOT NULL	question 구분을 위한 id
	qtype	smallint		question의 타입
	explain	text		question의 해설
	question_problem	varchar(255)		question의 지문
	right_answer	integer		question의 정답
	score	integer		question의 점수
	problem_id	varchar(255)		question이 속한 problem의 id
	example	text		question의 예시
	qetype	smallint		question의 예시의 타입
	created_at	timestamp(6) with time		question 생성일시
	updated_at	timestamp(6) with time		question 업데이트 일시
WRITE_ANSWER_BUNDLE	id	bigint	PRIMARY KEY	쓰기 정답 묶음 데이터 구분을 위한 id
	member	bigint	NOT NULL	쓰기 정답 묶음을 생성한 멤버의 id
	is_graduated	boolean	default false	채점 여부
	created_at	timestamp(6) with time		묶음 생성일시
	updated_at	timestamp(6) with time		묶음 업데이트 일시
WRITE_PROBLEM	id	bigint	PRIMARY KEY, NOT NULL	쓰기 문제를 구분하기 위한 id
	wtype	smallint		쓰기 문제 유형
	problem	varchar(255)		쓰기 문제 내용

	created_at	timestamp p(6) with time zone		쓰기 문제 생성일
	updated_at	timestamp p(6) with time zone		쓰기 문제 수정일
WRITE_QUESTION	id	bigint	PRIMARY KEY, NOT NULL	쓰기 질문을 구분 하기 위한 id
	question	varchar(255)		쓰기 질문 내용
	explain	text		쓰기 질문 설명
	score	int		쓰기 질문 점수
	problem	bigint	NOT NULL	쓰기 문제 id
	created_at	timestamp p(6) with time zone		질문 생성일시
	updated_at	timestamp p(6) with time zone		질문 업데이트 일 시
WRITE_ANSWER	id	bigint	PRIMARY KEY, NOT NULL	쓰기 답변을 구분 하기 위한 id
	answer	text		쓰기 답변 내용
	score	int		쓰기 답변 점수
	reason	text		쓰기 점수 이유
	question	bigint	NOT NULL	질문 참조에 대한 id
	member	bigint	NOT NULL	답변 작성한 멤버
	bundle	bigint		답변 묶음 정보
	created_at	timestamp p(6) with time zone		답변 생성일시
	updated_at	timestamp p(6) with time zone		답변 업데이트 일 시

표5. 데이터베이스 명세

3. 스토리지

- **Amazon S3**를 사용하여 문제 이미지 및 듣기 음성과 같은 **비정형 데이터**를 안전하게 저장하고 관리한다. 이를 통해 대용량 데이터의 효율적 보관이 가능하며, 필요 시 빠르게 불러올 수 있다.

3.2.6.3. 서비스 아키텍처

각 서비스들은 모두 Docker를 이용해 마이크로 서비스 아키텍처 형태로 운영되며 각각의 독립적인 작동을 보장한다.

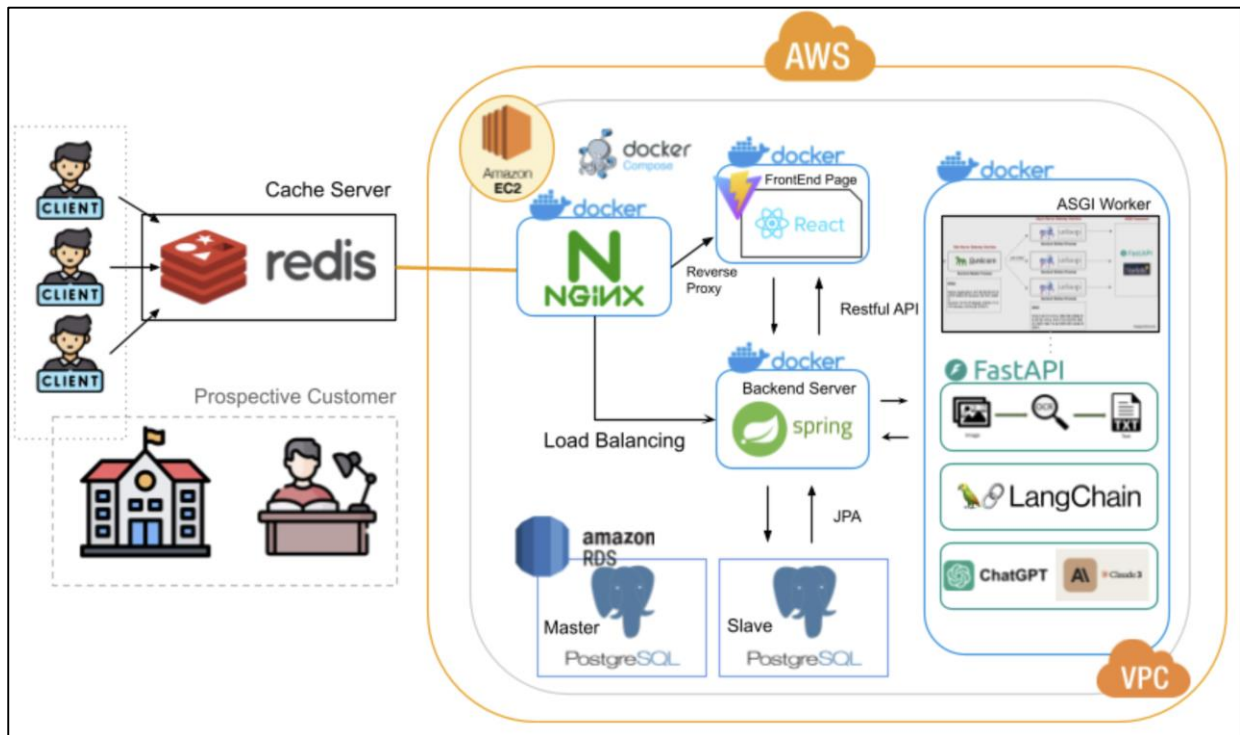


그림6. 전체 구조도

3.2.6.4. 서비스 시나리오

1. 사용자는 웹 브라우저를 통해 Vite + React JS로 구현된 CBT 랜딩 페이지에 접속 및 회원가입을 진행한다.
2. 비회원 및 인가되지 않은 사용자는 요청이 차단된다.
3. 로그인된 사용자의 요청은 NGINX를 통해 로드 밸런싱되어 적절한 Spring 백엔드 서버로 전달된다.
4. 백엔드 서버는 필요한 데이터를 PostgreSQL 데이터베이스 또는 Redis 캐시 서버에서 가져온다.

-
5. 비정형 데이터(이미지, 음성)는 Amazon S3에서 직접 로드된다.
 6. 생성 AI 기능이 필요한 경우, FastAPI를 통해 Python 서버의 Langchain 시스템과 통신하여 결과를 반환한다.
 7. 기본적으로 생성된 문제는 문제은행에 등록되어 재사용 가능하게 된다.
 8. 처리된 결과는 다시 사용자에게 전달되어 웹 페이지에 표시된다.

4. 연구 결과 분석 및 평가

4.1. 최종 완성 기능 목록

최종 결과보고서 작성 시점에서 완성된 기능들은 다음과 같다.

◆ 웹서비스

- ✓ 회원가입
- ✓ 메인 랜딩페이지
- ✓ 사용자 분류
 - 학생, 강사, 관리자
 - 사용자별 차등 권한 부여
- ✓ 사용자 시험 기록 및 통계 확인
- ✓ CBT (Computer Based Test)
 - 시험 구성 방식
 - ✧ 듣기, 읽기 토픽1,2 총 45유형
 - ✧ 기존 기출문제
 - ✧ AI 생성 시험
 - ✧ 랜덤 셔플(기출문제, AI 생성문제)
 - ✧ 원하는 유형만 선택하여 시험 커스터마이징
 - ✧ 사용자가 등록하는 특별 커스텀 시험
 - 시험 응시 및 채점 기능

✓ PBT

➤ 프린트가 가능한 PDF 생성 기능

✧ 종이 시험 커버, 정답지

✧ 1단, 2단 레이아웃

✓ 생성 문제 검수 및 수정 페이지

✓ 회원별 학습 통계 페이지

4.2. 성능 평가

자문 피드백 내용에 따라 백엔드 시스템에 **멀티 모듈 아키텍처**와 **batch 처리**를 도입하여 성능 지표를 정량적으로 분석하였다. 또한 문제 생성 방식을 한국어 선생님, 교수님들과 피드백하여 이를 통해 문제 생성의 정확도와 처리 속도를 개선하였다.

4.2.1. 생성 품질

다음은 구현된 웹 플랫폼에서 기존 기출문제와 생성된 문제의 비교이다.

<p>※ [1~3] ※ 무엇에 대한 내용입니까? <보기>와 같이 알맞은 것을 고르십시오.</p> <div style="border: 1px solid black; padding: 5px; text-align: center;"><p><보기></p><p>오늘은 월요일입니다. 내일은 화요일입니다.</p><p>① 공부 ② 얼굴 ③ 요일 ④ 계절</p></div> <p>1. (2점)</p> <div style="border: 1px solid black; padding: 5px;"><p>사과가 있습니다. 수박은 없습니다.</p></div> <p>① 시간 ② 과일 ③ 운동 ④ 가족</p> <p>2. (2점)</p> <div style="border: 1px solid black; padding: 5px;"><p>저는 열다섯 살입니다. 언니는 열일곱 살입니다.</p></div> <p>① 나이 ② 날짜 ③ 이름 ④ 나라</p> <p>3. (2점)</p> <div style="border: 1px solid black; padding: 5px;"><p>수미 씨는 회사원입니다. 컴퓨터 회사에 다닙니다.</p></div> <p>① 취미 ② 직업 ③ 약속 ④ 쇼핑</p> <p>※ [4~9] ※ <보기>와 같이 ()에 들어갈 말로 가장 알맞은 것을 고르십시오.</p> <div style="border: 1px solid black; padding: 5px; text-align: center;"><p><보기></p><p>단어를 모릅니다. ()을 찾습니다.</p><p>① 안경 ② 옷장 ③ 사전 ④ 지갑</p></div>	<p>※ [1~3] ※ 무엇에 대한 내용입니까? <보기>와 같이 알맞은 것을 고르십시오.</p> <div style="border: 1px solid black; padding: 5px; text-align: center;"><p><보기></p><p>선생님을 만납니다. 공부를 합니다.</p><p>① 학교 ② 요일 ③ 취미 ④ 날짜</p></div> <p>1. (2점)</p> <div style="border: 1px solid black; padding: 5px;"><p>어머니께서 요리하십니다. 오빠는 청소를 합니다.</p></div> <p>① 나이 ② 가족 ③ 음식 ④ 친구</p> <p>2. (2점)</p> <div style="border: 1px solid black; padding: 5px;"><p>지난 방학에 경주에 갔습니다. 올해는 제주도에 갈 것입니다.</p></div> <p>① 학교 ② 고향 ③ 교통 ④ 여행</p> <p>3. (2점)</p> <div style="border: 1px solid black; padding: 5px;"><p>그림을 그립니다. 음악을 듣습니다.</p></div> <p>① 노래 ② 취미 ③ 시험 ④ 운동</p> <p>※ [4~9] ※ <보기>와 같이 ()에 들어갈 말로 가장 알맞은 것을 고르십시오.</p> <div style="border: 1px solid black; padding: 5px; text-align: center;"><p><보기></p><p>숙제가 있습니다. 하지만 () 못 했습니다.</p><p>① 벌써 ② 아직 ③ 다시 ④ 늦게</p></div>
--	---

그림7. (좌) 23년 TOPIK I 읽기 시험 (우) 생성된 TOPIK I 읽기 시험

생성된 지문들이 원본 지문의 형태에서 벗어나지 않고 선택지 또한 원본과 비슷한 스타

일로 구성되어 있는 것을 볼 수 있다. 문제로서도 적절한 주제를 선정하여 선택지와 정답이 구성되어 있는 것으로 보인다. 짧은 문항 뿐 아니라 긴 지문에서의 비교를 보겠다.

<p>※ [31~32]※ 다음을 읽고 물음에 답하십시오.</p> <div> <p>저는 태국 친구하고 인주 역사 박물관에 갔습니다. (㉠) 친구는 한국어를 잘 못합니다. (㉡) 그런데 그 박물관에는 외국어로 설명해 주는 서비스가 있었습니다. (㉢) 친구는 태국어 설명 서비스를 신청했습니다. (㉣) 설명이 쉽고 재미있어서 친구가 정말 좋아했습니다.</p> </div> <p>31. 다음 문장이 들어갈 곳으로 가장 알맞은 것을 고르십시오. (2점)</p> <div> <p>잠시 후에 태국 사람이 와서 태국어로 역사를 설명해 줍니다.</p> <p>① ㉠ ② ㉡ ③ ㉢ ④ ㉣</p> </div> <p>32. 밑줄의 내용과 같은 것을 고르십시오. (3점)</p> <p>① 제 친구는 한국어를 잘합니다. ② 태국어 설명은 조금 어려웠습니다. ③ 제 친구는 태국어로 설명을 들었습니다. ④ 저는 외국어 설명 서비스를 신청했습니다.</p> <p>※ [33~34]※ 다음을 읽고 물음에 답하십시오.</p> <div> <p>지난주에 새 식당이 문을 열었습니다. 이 식당에는 여러 모양의 맛있는 가구들이 있습니다. 이 가구들은 모두 사장님이 만든 것인데 손님이 살 수도 있습니다. 저는 어제 동생과 그 식당에 처음 갔습니다. 우리는 꽃 모양 테이블에 앉아서 음식을 먹었습니다. 분위기도 좋고 음식도 맛있어서 우리는 다음에 또 (㉠).</p> </div> <p>33. ㉠에 들어갈 말로 가장 알맞은 것을 고르십시오. (2점)</p> <p>① 간 것 같습니다 ② 갈 수 없습니다 ③ 가면 안 됩니다 ④ 가기로 했습니다</p>	<p>※ [31~32]※ 다음을 읽고 물음에 답하십시오.</p> <div> <p>직원은 조직의 중요한 구성원입니다. (㉠) 그들은 회사의 목표를 달성하기 위해 다양한 역할을 수행합니다. (㉡) 직원의 만족도는 생산성과 직결되므로, 좋은 근무 환경을 제공하는 것이 필요합니다. (㉢) 교육과 개발 기회를 제공하면 직원의 역량이 향상됩니다. (㉣) 결국, 직원이 행복할 때 회사도 함께 성장할 수 있습니다.</p> </div> <p>31. 다음 문장이 들어갈 곳으로 가장 알맞은 것을 고르십시오. (2점)</p> <div> <p>또한, 직원의 의견을 존중하고 소통하는 문화가 중요합니다.</p> <p>① ㉠ ② ㉡ ③ ㉢ ④ ㉣</p> </div> <p>32. 밑글의 내용과 같은 것을 고르십시오. (3점)</p> <p>① 직원은 조직의 중요한 구성원입니다. ② 직원의 의견을 무시하고 소통하지 않는 문화가 중요합니다. ③ 교육과 개발 기회를 제공하면 직원의 역량이 감소됩니다. ④ 직원의 만족도는 생산성과 무관합니다.</p> <p>※ [33~34]※ 다음을 읽고 물음에 답하십시오.</p> <div> <p>합의는 두 사람 이상이 어떤 문제에 대해 의견을 일치시키는 과정을 말합니다. 합의는 대화와 협상을 통해 이루어지며, 서로의 입장을 이해하는 것이 중요합니다. 합의가 이루어지면, 모든 참여자는 그 결과에 대해 동의하게 됩니다. (㉠) 관계뿐만 아니라 비즈니스와 정치에서도 중요한 역할을 합니다. 합의는 갈등을 줄이고, 협력적인 관계를 형성하는 데 도움을 줍니다.</p> </div> <p>33. ㉠에 들어갈 알맞은 말을 고르십시오. (2점)</p> <p>① 개인적인 ② 기술적인 ③ 재정적인</p>
<div> <p>이런 상소문들을 저자는 항을 향한 개우미의 죽비 소리로 비유하고 있다.</p> <p>① ㉠ ② ㉡ ③ ㉢ ④ ㉣</p> </div> <p>※ [42~43]※ 다음을 읽고 물음에 답하십시오.</p> <div> <p>민욱이 주말에 친구 부부를 초대해도 되느냐고 물었을 때, 미연은 말없이 한참 뜰을 돌았다. 그녀는 그날 민욱과 함께 아파트의 발코니 벽을 페인트 칠할 계획이었다. (중략) 그들은 옥 개월 전에 E시의 아파트를 사서 이사했다. 지은 지 이십 년도 더 된 낡은 아파트였지만, 누구의 도움도 받지 않고 그들 스스로 이룬 일이었다. 미연은 직접 발뚎을 팔아 수도부터 세시까지 새로 손을 보았다. 하지만 집수리는 대강 되었고 해도, 미연이 생각해 둔 대로 인테리어가 완성되려면 아직 부족했다. “손님 초대는 좀 이르지 않아? 칠 단장도 될었는데……” “성재가 한국에 왔다.” 민욱이 무거운 목소리로 말했다. (중략) 민욱과 성재는 고등학교 동창이고, 그들을 통해 만난 미연과 연주도 십여 년이 넘는 인연이었다. 꽤나 가까운 사이였지만 성재가 사업 실패로 한국을 떠나면서 연락이 끊겼다. 미연은 자신이 연주의 이름을 오랫동안 잊고 있었다는 사실을 깨달았다. 더 이상 뭐라고 할 말이 없었다. (중략) 토요일 오전에 미연은 두 딸아이를 데리고 마트에 가서 고기와 채소를 샀다. 누군가를 초대해 음식을 대접하는 것이 정말 오랜만이었다.</p> </div> <p>42. 밑줄 친 부분에 나타난 ‘미연’의 심정으로 가장 알맞은 것을 고르십시오. (2점)</p> <p>① 후련하다 ② 불만스럽다 ③ 허전하다 ④ 자랑스럽다</p> <p>43. 밑글의 내용으로 알 수 있는 것을 고르십시오. (2점)</p> <p>① 미연은 민욱과 성재를 통해 연주를 알게 되었다. ② 민욱은 주변 도움을 받아 E시의 새 아파트를 샀다. ③ 미연은 옥 개월 전에 산 아파트로 이사 갈 계획이다. ④ 민욱은 인테리어를 완성한 후에 마트에 장을 보러 갔다.</p> <p>※ [44~45]※ 다음을 읽고 물음에 답하십시오.</p> <div> <p>‘수로왕 신화’는 왕이 앞에서 탄생하는 여느 건국 신화와 유사하다. 고대에 아직 나라가 없던 낙동강 하류에 황금알 여섯 개가 하늘에서 내려왔다. 그중 맨저 알을 깨고 나온 아</p> </div>	<div> <p>우리는 기쁨, 슬픔, 분노, 사랑 등 다양한 감정을 경험하며 살아간다.</p> <p>① ㉠ ② ㉡ ③ ㉢ ④ ㉣</p> </div> <p>※ [42~43]※ 다음을 읽고 물음에 답하십시오.</p> <div> <p>‘한적한 시골 마을의 정류장에서, 지민은 버스를 기다리고 있었다. 매일 같은 노선이었으며 이미 자신이 몇 번째로 기다리는지를 알고 있었다. 버스가 언제 오려나? 지민은 불안한 표정으로 중얼거렸다. 그때, 그녀의 친구 아라가 나타났다. 너 또 그렇게 기다리기만 할 거야? 아라는 지민의 어깨를 툭 쳤다. 지민은 체중 없이 다가오는 버스를 기다리는 게 이렇게 힘든 일인지 몰랐어.라고 대답했다. 버스가 도착하는 순간, 지민의 가슴은 덩달아 설렘했다. 드디어 왔네! 아라가 미소 지으며 말했다. 그러나 두 사람은 버스에 오르자마자 떠오르는 불안감에 짓눌렸다. 이 버스로 정류장마다 서는 거야? 지민이 물었다. 그렇지, 그리고 우리는 또 많은 사람들과 함께 이 길을 가야 해. 아라가 대답하며 고개를 끄덕였다. 그들은 서로의 눈빛에서 버스라는 낯선 공간에 대한 두려움을 읽었다. 지민은 나는 할랄 이 버스를 타면 겁이 나,라며 마음 속의 불안을 토로했다. 아라는 하지만 새로운 곳에 가는 것도 기대되지 않아?라고 말하며 지민에게 용기를 주었다. 버스가 움직이기 시작하자, 두 사람의 얼굴에도 미소가 떠올랐다. 이제 한번 신나게 떠나보자! 아라가 외쳤다. 그리고 기어코, 그들은 새로운 시작을 향해 나아갔다.’</p> </div> <p>42. 밑줄 친 부분에 나타난 ‘지민’의 심정으로 가장 알맞은 것을 고르십시오. (2점)</p> <p>① 두려운 ② 무관심한 ③ 안정된 ④ 기쁜</p> <p>43. 밑글의 내용으로 알 수 있는 것을 고르십시오. (2점)</p> <p>① 지민은 한적한 시골 마을의 정류장에서 버스를 기다리고 있었다. ② 아라는 지민과 함께 버스를 기다리지 않았다. ③ 버스가 도착하기 전, 지민은 친구와 대화를 나누지 않았다. ④ 지민은 매일 다른 노선의 버스를 기다리고 있었다.</p> <p>※ [44~45]※ 다음을 읽고 물음에 답하십시오.</p> <div> <p>이름과 생일을 기록한(1)은 자신의 이름을 알리고, 몸집과 수업을 계획한(2)은 학생들의 자</p> </div>

그림8. (좌) 23년 TOPIK II 읽기 시험 (우) 생성된 TOPIK II 읽기 시험

문장의 구조가 살짝 경직되어 있긴 하지만 적당한 길이의 지문부터 매우 긴 지문에서도 원 시험 형태의 스타일을 간직하고 있는 것을 볼 수 있다. 또한 긴 지문에서도 선택지가 잘 구성되어 있는 것을 볼 수 있다.

4.2.2. 생성 과정

```

연휴
gpt-4o-mini |
휴가

OpenAI Tokens Used: 58
  Prompt Tokens: 56
  Completion Tokens: 2
Successful Requests: 1
Total Cost (USD): $4e-05
Total Cost (KRW): ₩0.054

[{'문장': '차는 빨간색입니다. 버스는 노랑색입니다.', '주제어': '도로'}, {'문장': '사상은 중요하다. 사상은 변할 수 있다.', '주제어': '사상'}]
gpt-4o-mini |
휴가는 즐겁다.
휴가는 필요하다.

OpenAI Tokens Used: 120
  Prompt Tokens: 108
  Completion Tokens: 12
Successful Requests: 1
Total Cost (USD): $9e-05
Total Cost (KRW): ₩0.1215

gpt-4o-mini |
휴식

OpenAI Tokens Used: 76
  Prompt Tokens: 74
  Completion Tokens: 2
Successful Requests: 1
Total Cost (USD): $6e-05
Total Cost (KRW): ₩0.081

gpt-4o-mini |
1. 비행
2. 전쟁
3. 농사

OpenAI Tokens Used: 66
  Prompt Tokens: 51
  Completion Tokens: 15
Successful Requests: 1
Total Cost (USD): $5e-05
Total Cost (KRW): ₩0.0675

gpt-4o |
answer: 휴식
explain: '휴식'은 '휴가'와 의미가 가장 유사하며, 문맥상 '휴식'이 '휴가'와 잘 어울립니다. '농사', '비행', '전쟁'은 문맥상 어울리지 않습니다.

OpenAI Tokens Used: 153
  Prompt Tokens: 89
  Completion Tokens: 64
Successful Requests: 1
Total Cost (USD): $0.00306
Total Cost (KRW): ₩4.131

{'example': '휴가는 즐겁다. 휴가는 필요하다.',
'selector': ['농사', '휴식', '비행', '전쟁'],
'answer': '휴식',
'eval_answer': '휴식',
'eval_explain': '"휴식'은 '휴가'와 의미가 가장 유사하며, 문맥상 '휴식'이 '휴가'와 잘 어울립니다. '농사', '비행', '전쟁'은 문맥상 어울리지 않습니다.",
'total_cost': 4.456}

```

그림9. TOPIK I 읽기 유형 1번 어휘 선택 문제 생성과정

쿼리가 없는 문제는 주제가 자동 설정되게 되어있다. 연휴라는 단어가 선정되었고 필터

링 된 교육적 주제로 휴가를 제시하였다. RAG를 통해 비슷한 유형의 문장을 가져오고 이를 통해 새로운 지문을 구성한다. 이후 선택지를 구성하고 결과 검증 이후 마지막 example(지문), selector(선택지), answer(정답), eval_answer(검증 정답), eval_explain(해설)을 출력한다. LLM의 비용 추적을 구현해 해당 문제 생성 비용이 4.45₩임 또한 알 수 있다.

4.2.1. 생성 비용

```
from datetime import datetime
import glob

folder_path = '/home/shinbg/topik-korea-data/data/problem_data'

# 'generate_exam'로 시작하고 '.json'으로 끝나는 모든 파일을 검색합니다.
file_pattern = os.path.join(folder_path, "generate_exam_*.json")
files = glob.glob(file_pattern)

if not files:
    print("JSON 파일을 찾을 수 없습니다.")
else:
    Exam_cost_list = []
    for file in files:
        with open(file, "r") as json_file:
            exam_cost = []
            json_data = json.load(json_file)
            json_data = example_preprocess(json_data)
            for item in json_data:
                for question in item:
                    exam_cost.append(question['total_cost'])
            print("Exam Total Cost: ", sum(exam_cost))
            print("Problem Avg Cost: ", sum(exam_cost)/len(exam_cost))
            Exam_cost_list.append(sum(exam_cost))
    print("Total TESTSET Cost: ", sum(Exam_cost_list))
    print("Total TESTSET Avg Cost: ", sum(Exam_cost_list)/len(Exam_cost_list))
```

Exam Total Cost: 396.23849999999993
 Problem Avg Cost: 13.207949999999999
 Exam Total Cost: 381.48299999999999
 Problem Avg Cost: 12.716099999999996
 Exam Total Cost: 387.93599999999999
 Problem Avg Cost: 12.931199999999997
 Exam Total Cost: 411.07499999999993
 Problem Avg Cost: 13.702499999999997
 Exam Total Cost: 400.14000000000004
 Problem Avg Cost: 13.338000000000001
 Exam Total Cost: 358.50600000000003
 Problem Avg Cost: 11.9502
 Exam Total Cost: 420.28199999999999
 Problem Avg Cost: 14.009399999999998
 Exam Total Cost: 392.067
 Problem Avg Cost: 13.068900000000001
 Exam Total Cost: 374.04450000000001
 Problem Avg Cost: 12.468150000000003
 Exam Total Cost: 386.451
 Problem Avg Cost: 12.8817
 Total TESTSET Cost: 3908.223
 Total TESTSET Avg Cost: 390.8223

그림10. 생성 비용 추적 및 계산 코드

읽기 I 유형 기준 10개 시험(각 40문제) 생성비용을 추적해보았다. 각 문제의 평균 생성 비용은 13W 정도가 발생하며 40문제로 이루어진 시험지 1개를 생성하면 400W 정도가 사용되는 것을 알 수 있다. 해당 통계 값은 GPT-3.5-Turbo모델을 사용했던 값으로 현재는 30% 더 저렴한 GPT-4o-mini를 사용하고 있으며 비용을 계속 최적화 하고 있다.

4.2.2. 생성 속도

본 과제에서는 생성 유형 독립, 각 유형 간의 비동기 생성, 각 문제의 생성에 멀티프로세싱을 동원하여 시험 생성 단계에서 병렬 처리를 구현하였다.

```

• from Problem_Type_Template.reading_1_problem_type_8_2 import generate_reading_1_problem_type_8_2

result = generate_reading_1_problem_type_8_2(default_model="gpt-4o-mini",verbose=True)

display(result)
✓ 16.4s

```

```

(topik) (base) shinnbg@shinnbg-ws:~/topik-korea-data$ /home/shinnbg/anaconda3/envs/topik/bin/python /home/shinnbg/topik-korea-data/exam_generator.py
Generating reading_1_problem_type_2: 100%| 3/3 [00:12:00:00, 4.31s/it]
Generating reading_1_problem_type_1_2: 100%| 1/1 [00:23:00:00, 23.43s/it]
Generating reading_1_problem_type_3: 100%| 3/3 [00:23:00:00, 7.98s/it]
Generating reading_1_problem_type_1_2: 100%| 2/2 [00:24:00:00, 12.14s/it]
Generating reading_1_problem_type_1_2: 100%| 1/1 [00:24:00:00, 24.57s/it]
Generating reading_1_problem_type_1_1: 100%| 3/3 [00:24:00:00, 8.21s/it]
Generating reading_1_problem_type_1_2: 100%| 1/1 [00:24:00:00, 24.84s/it]
Generating reading_1_problem_type_1_2: 100%| 1/1 [00:25:00:00, 25.13s/it]
Generating reading_1_problem_type_7: 100%| 1/1 [00:03:00:00, 3.29s/it]
Generating reading_1_problem_type_6: 100%| 2/2 [00:07:00:00, 3.94s/it]
Generating reading_1_problem_type_8_1: 100%| 1/1 [00:11:00:00, 11.12s/it]
Generating reading_1_problem_type_4: 100%| 3/3 [00:24:00:00, 8.16s/it]
Generating reading_1_problem_type_5: 100%| 4/4 [00:38:00:00, 9.59s/it]
Generating reading_1_problem_type_8_2: 100%| 4/4 [00:39:00:00, 9.75s/it]
Generating Problems: 100%| 14/14 [01:04:00:00, 4.63s/it]

```

그림11. (상) 읽기 I 유형 8 생성 시간 (하) 병렬처리 시험 생성기 생성 시간

다음은 구현된 병렬 처리의 성능을 보여준다. 일반적으로 한 문제의 동기적 생성은 16초 정도가 걸렸지만, 구현한 병렬 처리를 통해 8개 유형, 하위 문제는 평균 5개를 가지고 총 40문항의 시험을 생성한 결과 1분 4초(64초) 정도가 걸렸다. 만일 동기적으로 진행했다면 $8 \times 5 \times 16 = 640$ (약 10분) 가량이 걸리는 것으로 10배 가량 생성 속도를 대폭 단축하였다.

4.3. 결과물

과제 결과물은 **모의시험 플랫폼**의 형태로 완성되었으며, 다양한 문제 유형을 자동 생성할 수 있는 시스템이 구축되었다. 학습자는 자신만의 시험지를 구성하여 실전 대비 학습을 할 수 있으며, 자동 채점 및 학습 통계 기능을 통해 학습 효율성을 크게 향상시킬 수 있다.

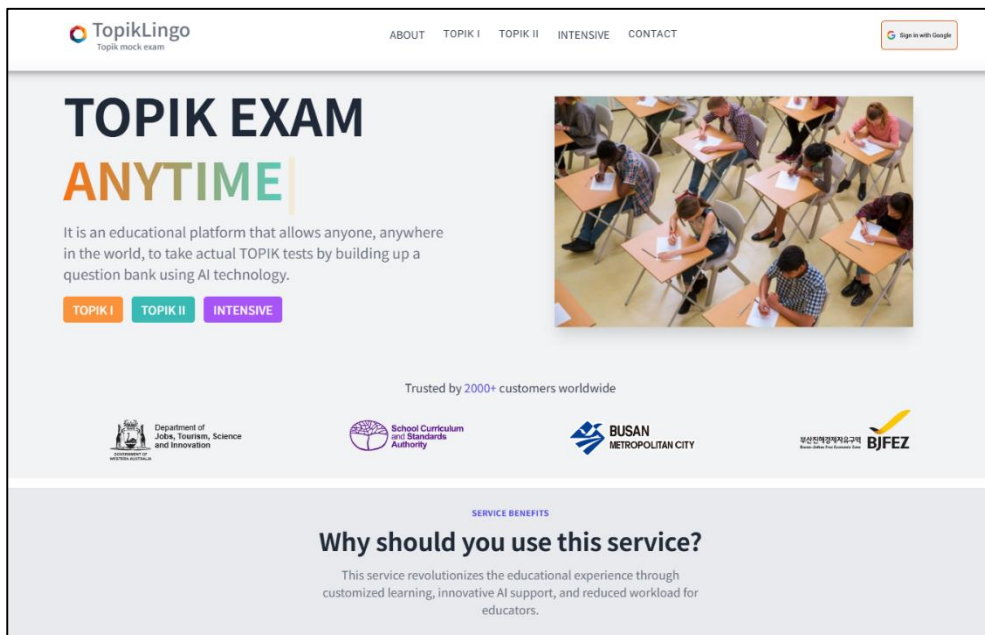


그림12. 메인 랜딩 페이지

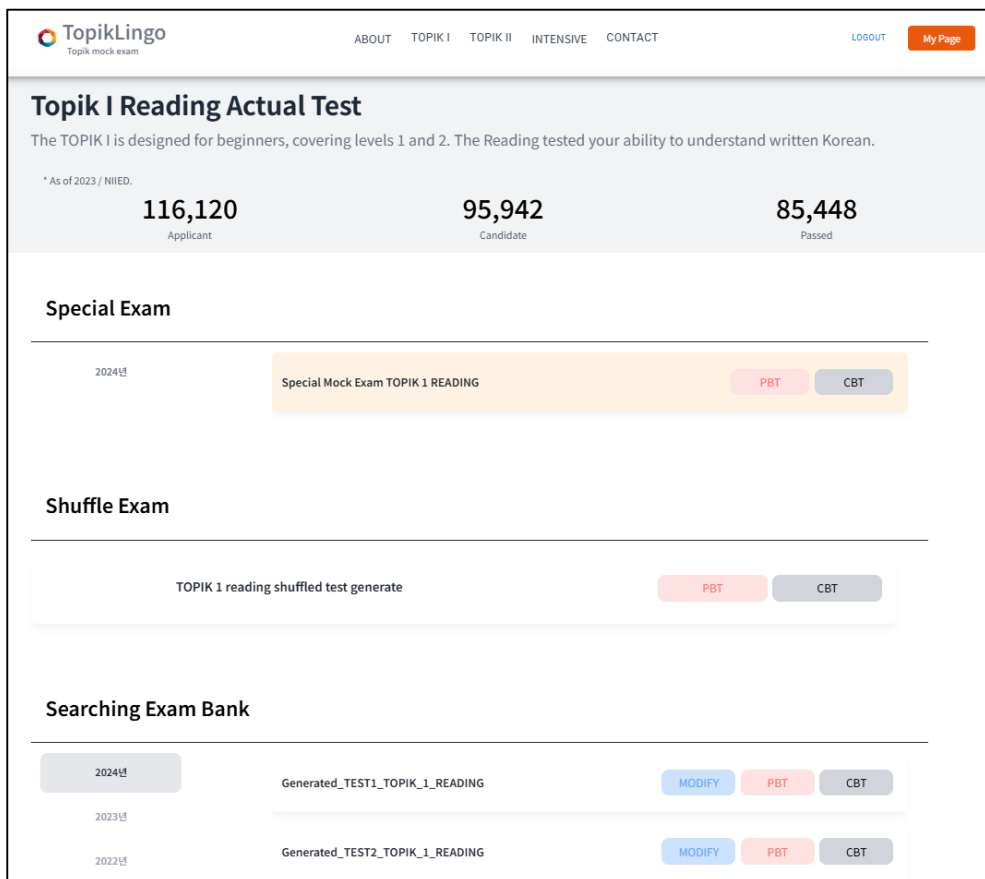


그림13. 시험 선택 페이지(특별시험, 문제 셔플, 문제 은행 조회)

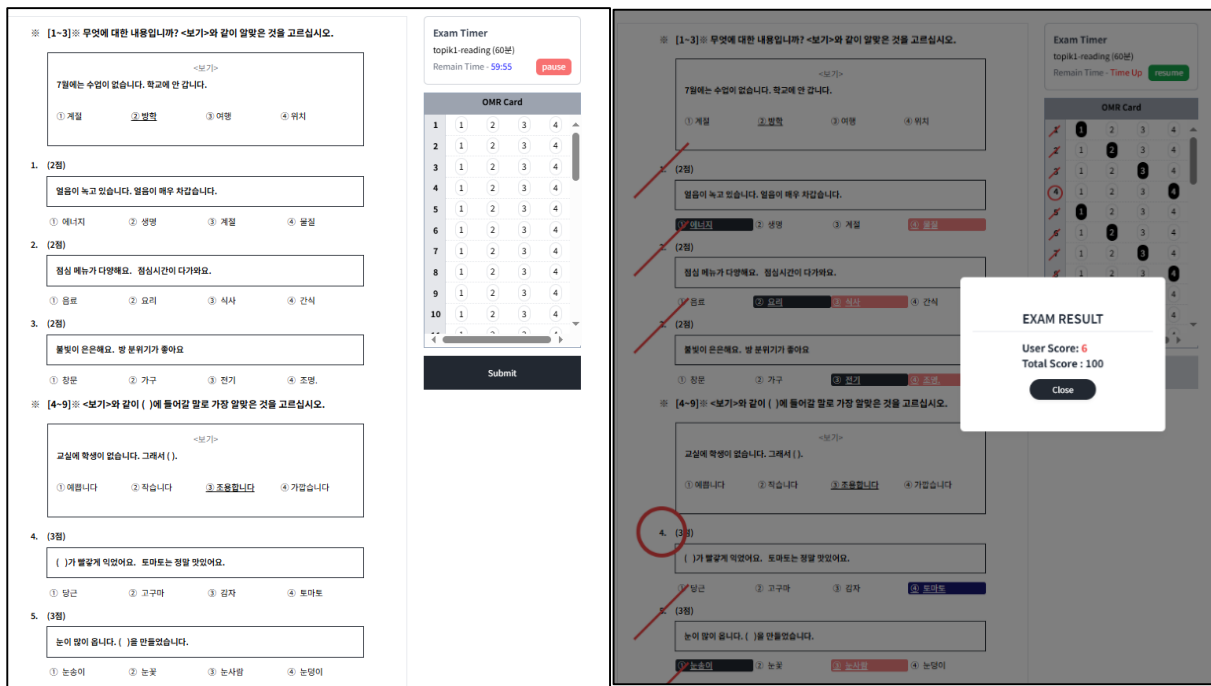


그림14. 읽기 CBT(시험지, 타이머, OMR카드, 채점)

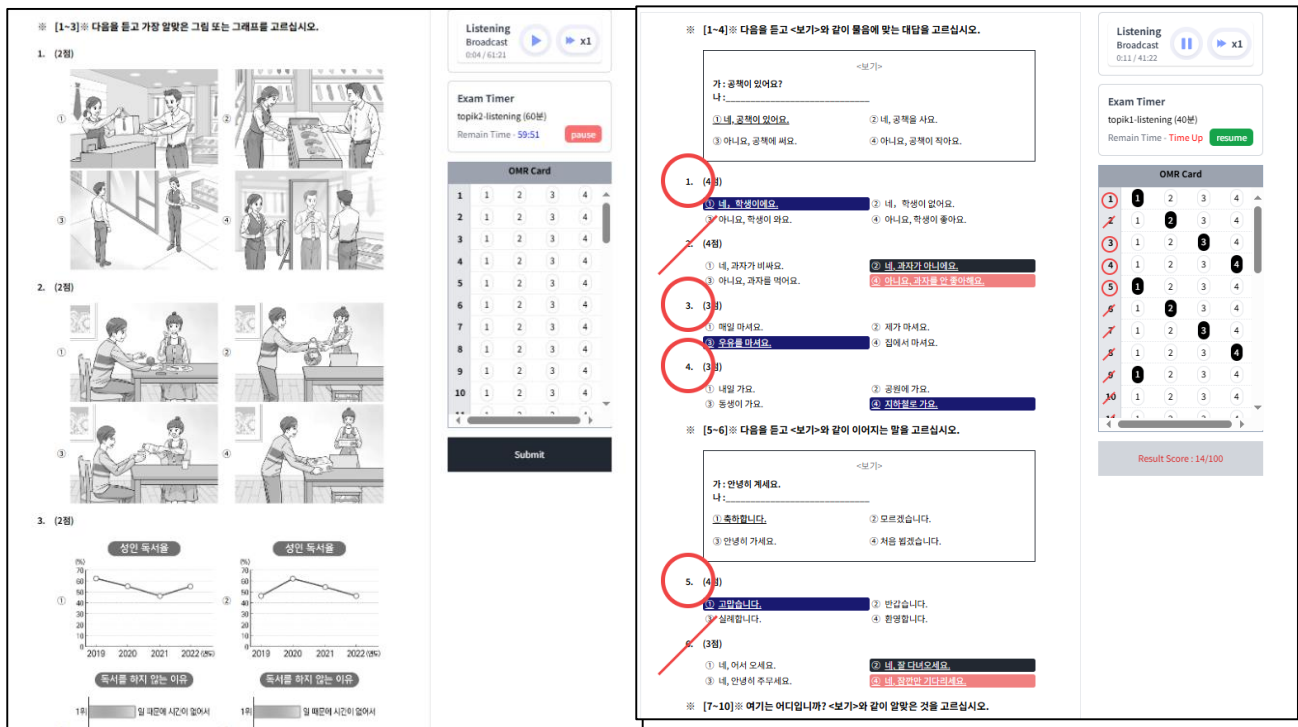


그림15. 듣기 CBT(듣기 방송 재생, 시험지, 타이머, OMR카드, 채점)

TopikLingo
Topik mock exam

ABOUT TOPIK I TOPIK II INTENSIVE CONTACT

LOGOUT My Page

* As of 2023 / NIIED.

418,001 Applicant 333,289 Candidate 241,587 Passed

INTENSIVE EXAM GENERATOR

TOPIK1 Reading

TOPIK2 Reading

문법

유의 문법/표현

광고/소재

도표/그래프

순서 배열

빈칸 1

빈칸 2

실제/알은 것

제목 찾기

내용 일치

주제 찾기

문장 위치

문학

비문학

정보

종합

Selected Type

TOPIK1 Reading

어휘

그림/표

빈칸 1

순서 나열

TOPIK2 Reading

광고/소재

도표/그래프

Selected Type Num

Count

* Integer range for 1-5

Test Setting

Total Selected Type Num

6

Exam Title

☒ Include Cover

☒ Include Answer

☒ Print Layout

☐ 1 Layer (Experimental) ☒ 2 Layer (Default)

PBT CBT

그림16. 타입 선택 시험(원하는 종류의 시험만 선택하여 풀기)

TOPIK

제 2024년
TOPIK
Actual Test

The 2024 Actual Test of Proficiency in Korean

ExamID: 46c78928-3871-11ef-af51-e0d55eba17b3

TOPIK I
READING

Total 40 Questions

Time Limit: 60 minutes

TheGrapeBine

46c78928-3871-11ef-af51-e0d55eba17b3

※ [1~3]은 무엇에 대한 내용입니까? <보기>
와 같이 알맞은 것을 고르십시오.

<보기>
7월에는 수확이 없습니다. 학교에 간 겁니다.

① 계절 ② 보람 ③ 여행 ④ 위치

1. (2점)
방문이 늦고 있습니다. 방문이 매우 지체됩니다.

① 늦다 ② 생략 ③ 계정 ④ 물품

2. (2점)
탐험 계획이 다양해요. 탐험시간이 다가요.

① 유물 ② 요리 ③ 역사 ④ 간식

3. (2점)
통일이 큰데요. 한 문자가 좋아요

① 방문 ② 가구 ③ 잔기 ④ 조형.

※ [4~9]은 <보기>와 같이 ()에 들어갈 말로
가장 알맞은 것을 고르십시오.

<보기>
교실에 학생이 있습니다. 그래서 ().

① 애통하다 ② 걱정하다 ③ 기쁘다
다

4. (3점)
()가 별장게 먹었어요. 통이들은 정말 맛있어요.

① 닭 ② 고구마 ③ 감자 ④ 토마토

5. (3점)
눈이 많이 옵니다. ()를 만들었습니다.

① 눈송이 ② 눈꽃 ③ 눈사람 ④ 눈덩이

6. (2점)
문자 메시지 ()를 보내는 중이에요.

① 발송함 ② 수감함 ③ 발송함 ④ 전달함

① 여로 ② 여로 ③ 여로 ④ 여로

7. (2점)
당근은 주황색 ()로 보충하는 식품입니다.

① 이다 ② 아니다 ③ 존재함 ④ 없다

8. (2점)
눈물 () 멈추지 않습니다. 마음이 너무 아픕니다.

① 한 ② 가 ③ 며 ④ 는

9. (2점)
강물은 저장 수 있습니다. 대개 ()해결해 주세요.

① 으르 ② 세어 ③ 보 ④ 세어

TOPIK

Answer Sheet

ExamID: 46c78928-3871-11ef-af51-e0d55eba17b3

Question	Answer
1	(A) 번
2	(B) 번
3	(A) 번
4	(A) 번
5	(B) 번
6	(B) 번
7	(A) 번
8	(B) 번
9	(B) 번
10	(A) 번
11	(A) 번
12	(A) 번
13	(A) 번
14	(A) 번
15	(A) 번
16	(B) 번
17	(A) 번
18	(A) 번
19	(A) 번

그림17. PBT 변환(커버, 시험지, 정답지 형태로 프린트할수 있는 PDF)

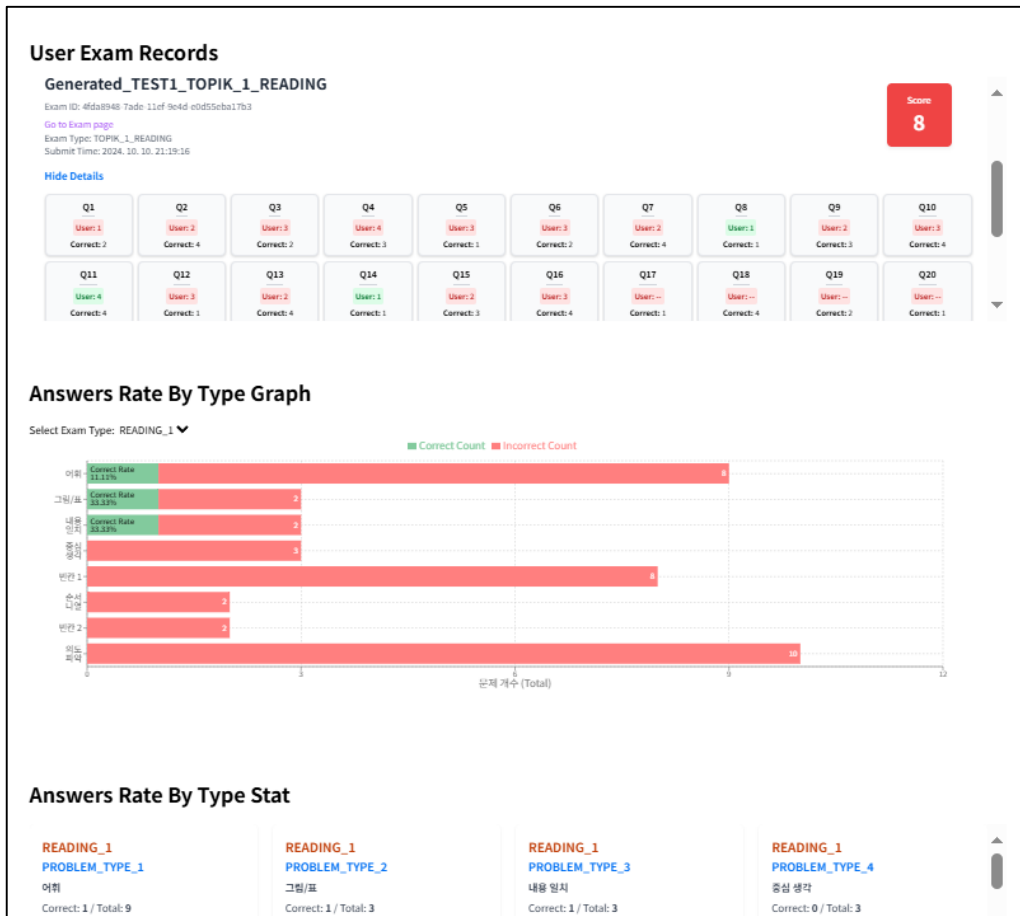


그림18. 유저 페이지(상세정보, 시험 기록, 타입별 정답률 통계)

※ [1~3] ※ 무엇에 대한 내용입니까? <보기>와 같이 알맞은 것을 고르십시오.

<보기>

7월에는 수확이 없습니다. 학교에 간 겁니다.

① 계절 ② 문화 ③ 여행 ④ 위치

1. (2점)

열음이 높고 있습니다. 열음이 매우 차갑습니다.

열음이 높고 있습니다. 열음이 매우 차갑습니다.

① 에너지 ② 생명 ③ 계절 ④ 물질

에너지 생명 계절 물질

2. (2점)

참심 매뉴가 다양해요. 참심시간이 다가워요.

참심 매뉴가 다양해요. 참심시간이 다가워요.

① 음료 ② 요리 ③ 식사 ④ 간식

음료 요리 식사 간식

3. (2점)

불빛이 은은해요. 병 분위기가 좋아요

불빛이 은은해요. 병 분위기가 좋아요

① 향문 ② 가구 ③ 전기 ④ 조명

향문 가구 전기 조명

입업 활동 지원봉사 운동 직무

※ [9~10] ※ 다음 글 또는 그래프의 내용과 같은 것을 고르십시오.

9. (2점)

사람들은 사람들의 일상생활에서 중요한 역할을 한다. 다양한 상품을 판매하며, 고객은 필요한 물건을 쉽게 구입할 수 있다. 특히 동네 상점은 지역 주민들과의 소통 공간이 되기도 하며, 상점 주인과의 친밀한 관계를 통해 지역 사회가 더욱 온화해진다. 요즘은 온라인 쇼핑이 대세지만, 많은 사람들은 상점에서 직접 물건을 보고 구매하는 즐거움을 여전히 선호한다.

상점은 사람들의 일상생활에서 중요한 역할을 한다. 다양한 상품을 판매하며, 고객은 필요한 물건을 쉽게 구입할 수 있다. 특히 동네 상점은 지역 주민들과의 소통 공간이 되기도 하며, 상점 주인과의 친밀한 관계를 통해 지역 사회가 더욱 온화해진다. 요즘은 온라인 쇼핑이 대세지만, 많은 사람들은 상점에서 직접 물건을 보고 구매하는 즐거움을 여전히 선호한다.

① 상점에서는 고객이 필요한 모든 것을 무료로 제공한다.
상점에서는 고객이 필요한 모든 것을 무료로 제공한다.

② 상점은 온라인 쇼핑보다 항상 더 비쌀 것이다.
상점은 온라인 쇼핑보다 항상 더 비쌀 것이다.

③ 상점은 지역 사회의 소통 공간 역할을 한다.
상점은 지역 사회의 소통 공간 역할을 한다.

④ 모든 사람들이 상점에서만 물건을 산다.
모든 사람들이 상점에서만 물건을 산다.

10. (2점)

게스차는 의사소통에서 중요한 역할을 한다. 사람들은 말뿐만 아니라 손이나 몸을 사용하여 감정이나 의도를 전달한다. 예를 들어, 손을 흔드는 것은 인사하는 게스차이며, 엄지손가락을 올리는 것은 칭찬이나 승인을 나타낸다. 연구에 따르면, 게스차는 언어적 표현보다 더 강력한 영향을 미칠 수 있으며, 문화에 따라 그 의미가 달라질 수 있다. 따라서 다양한 게스차를 이해하는 것은 효과적인 소통을 위해 필수적이다.

게스차는 의사소통에서 중요한 역할을 한다. 사람들은 말뿐만 아니라 손이나 몸을 사용하여 감정이나 의도를 전달한다. 예를 들어, 손을 흔드는 것은 인사하는 게스차이며, 엄지손가락을 올리는 것은 칭찬이나 승인을 나타낸다. 연구에 따르면, 게스차는 언어적 표현보다 더 강력한 영향을 미칠 수 있으며, 문화에 따라 그 의미가 달라질 수 있다. 따라서 다양한 게스차를 이해하는 것은 효과적인 소통을 위해 필수적이다.

① 게스차는 언제나 동일한 의미를 가진다.
게스차는 언제나 동일한 의미를 가진다.

② 게스차는 언제나 동일한 의미를 가진다.
게스차는 언제나 동일한 의미를 가진다.

③ 게스차는 의사소통에서 중요하지 않다.
게스차는 의사소통에서 중요하지 않다.

알려지는 음악의 맥락을 결정짓는 핵심 요소이다.

④ 멜로디는 음악의 감정을 전달하는 중요한 요소이다.
멜로디는 음악의 감정을 전달하는 중요한 요소이다.

※ [46~47] ※ 다음을 읽고 물음에 답하십시오.

제품은 현대 사회에서 필수적인 요소로 자리 잡고 있다. 소비자들은 다양한 선택지 속에서 품질과 가성비를 동시에 고려하게 된다. 따라서, 기업은 지속적인 연구 개발과 혁신을 통해 시장의 요구에 부응해야 한다. 또한, 사회적 책임을 다하는 제품 생산이 점점 더 중요해지고 있다. 예를 들어, 친환경 제품이나 공정 거래 제품은 소비자에게 긍정적인 이미지를 준다. 결국, 성공적인 제품은 소비자와의 신뢰 관계를 구축하는 데 중점을 두어야 한다. 이러한 요소들이 결합될 때, 제품은 단순한 물건 이상의 가치를 지니게 된다.

제품은 현대 사회에서 필수적인 요소로 자리 잡고 있다. 소비자들은 다양한 선택지 속에서 품질과 가성비를 동시에 고려하게 된다. 따라서, 기업은 지속적인 연구 개발과 혁신을 통해 시장의 요구에 부응해야 한다. 또한, 사회적 책임을 다하는 제품 생산이 점점 더 중요해지고 있다. 예를 들어, 친환경 제품이나 공정 거래 제품은 소비자에게 긍정적인 이미지를 준다. 결국, 성공적인 제품은 소비자와의 신뢰 관계를 구축하는 데 중점을 두어야 한다. 이러한 요소들이 결합될 때, 제품은 단순한 물건 이상의 가치를 지니게 된다.

46. 밑줄에 나타난 밑자의 태도로 가장 알맞은 것을 고르십시오. (2점)

① 소비자들은 품질보다 가격만을 중시하고 있다고 주장하고 있다.
소비자들은 품질보다 가격만을 중시하고 있다고 주장하고 있다.

② 친환경 제품의 중요성을 간과하고 있다고 언급하고 있다.
친환경 제품의 중요성을 간과하고 있다고 언급하고 있다.

③ 제품의 사회적 책임은 기업에 큰 부담이 된다고 생각하고 있다.
제품의 사회적 책임은 기업에 큰 부담이 된다고 생각하고 있다.

④ 기업은 소비자와의 신뢰 관계를 구축하는 것이 중요하다고 강조하고 있다.
기업은 소비자와의 신뢰 관계를 구축하는 것이 중요하다고 강조하고 있다.

47. 밑줄의 내용과 같은 것을 고르십시오. (2점)

① 친환경 제품은 소비자에게 부정적인 이미지를 만든다.
친환경 제품은 소비자에게 부정적인 이미지를 만든다.

② 제품속 현대 사회에서 소비자의 신뢰 관계를 구축하는 데 중점을 두어야 한다.
제품은 현대 사회에서 소비자의 신뢰 관계를 구축하는 데 중점을 두어야 한다.

③ 기업은 연구 개발을 통해 시장의 요구를 무시해야 한다.
기업은 연구 개발을 통해 시장의 요구를 무시해야 한다.

④ 소비자들은 품질만 고려하여 제품을 선택한다.
소비자들은 품질만 고려하여 제품을 선택한다.

그림19. 시험 검수 페이지(정답은 검은색으로 표시, 모든 글귀 수정 가능)

4.4. 테스트

현재 실제 학원을 운영하는 선생님들과 함께 플랫폼에 대한 검증과 실증을 진행 중이다.

2024년 9월 26일 인제대에서 본 과제의 플랫폼으로 파일럿 테스트를 진행하였다.



그림 20. 플랫폼 파일럿 테스트

외국인 유학생 26명을 대상으로 진행하였는데, 가입의 편의를 위해 구글 로그인 방식으로 구현했음에도 생각보다 외국인들이 구글 아이디를 많이 사용하고 있지 않아 모두가 로그인을 해서 테스트 시험을 진행해보는 데 30분이 걸렸다. 또한 현재 튜토리얼 안내 등이 없어 따로 감독이 필요하고, 컴퓨터를 조금 다룰 줄 알아야 하는 것이 문제로 꼽혔다. 다행히 플랫폼의 전반적인 모든 기능은 잘 작동하였고 시험에 관련해서는 불편 사항이 없었다. 전반적인 관리를 위해 성적 데이터, 시험 시도 기록 조회 API를 개발하여 학생들의 학습 상황을 모니터링 했으며 테스트 날 26명의 학습 데이터가 잘 확인되었다.

Save your result: response_172862521.csv Download Result EOL: [CRLF]

department	name	email	examName	score	date
Inje University	이준호 (Joonho Lee)	joonho.lee@inje.ac.kr	Special Mock Exam TOPIK 2 READING	48	2024-09-26T14:50:18.194471
Inje University	김민준 (Minjun Kim)	minjun.kim@inje.ac.kr	Special Mock Exam TOPIK 2 READING	60	2024-09-26T14:46:05.154556
Inje University	박지민 (Jimin Park)	jimmin.park@inje.ac.kr	Special Mock Exam TOPIK 2 READING	64	2024-09-26T14:48:19.576377
Inje University	이성민 (Seungmin Lee)	seungmin.lee@inje.ac.kr	Special Mock Exam TOPIK 2 READING	36	2024-09-26T14:39:06.797052
Inje University	김민준 (Minjun Kim)	minjun.kim@inje.ac.kr	Special Mock Exam TOPIK 2 READING	80	2024-09-26T14:43:03.635515
Inje University	이준호 (Joonho Lee)	joonho.lee@inje.ac.kr	Special Mock Exam TOPIK 2 READING	50	2024-09-26T14:39:19.521764
Inje University	김민준 (Minjun Kim)	minjun.kim@inje.ac.kr	Special Mock Exam TOPIK 2 READING	72	2024-09-26T14:41:37.604888
Inje University	이성민 (Seungmin Lee)	seungmin.lee@inje.ac.kr	Special Mock Exam TOPIK 2 READING	66	2024-09-26T14:40:25.143538
Inje University	김민준 (Minjun Kim)	minjun.kim@inje.ac.kr	Special Mock Exam TOPIK 2 READING	36	2024-09-26T14:25:22.586237
Inje University	이성민 (Seungmin Lee)	seungmin.lee@inje.ac.kr	Special Mock Exam TOPIK 2 READING	68	2024-09-26T14:45:01.199446

그림 21. 플랫폼 파일럿 테스트 결과

5. 결론 및 향후 연구 방향

이번 과제를 통해 AI 기반 문제 자동 생성 및 모의시험 플랫폼의 가능성을 확인하였다. 충분히 생성AI를 이용하여 교육적 문제를 다루고 실제 활용 가능한 수준의 문제 생성이 가능하다고 생각되는데, 해당 생성 결과를 실제 TOPIK 문제를 내는 자문위원 교수님께 여쭙보았고 엄격한 기준이 있는 빈칸 문제를 제외하고 지문의 퀄리티는 충분히 활용 가능할거라 하셨다. 이러한 부분은 추가적인 프롬프트 튜닝과 형태소 분석을 통한 자연어 처리로 좀 더 세밀하게 처리해 고도화 할 수 있는 여지가 많이 남았다고 생각된다.

다만, 현재 사용중인 생성AI가 OpenAI에 의존적이다 보니 LLM Response를 Parsing하는 과정에서 OpenAI가 모델을 새로운 버전으로 업데이트 하면 기존의 결과 제어 프롬프트의 결과가 달라져 Parsing이 되지 않고 에러를 뿜는 경향이 있다. 현재는 이러한 불안정성으로 문제 생성기가 API로 Serving은 되어 있지만 웹사이트에서 사용자가 실시간으로 생성 후 탑재하기엔 오류의 확률이 높아 관리자가 수동으로 확인 이후 문제 은행에 등록할 수 있도록 되어있다.

생성기와는 별개로, 다행히 플랫폼의 완성도는 파일럿 테스트에서 큰 문제가 없었고 시험UI 대해선 호평이었지만 UX측면에서 안내 튜토리얼, 더욱 간편한 디자인, 모바일 화면에서의 시험 환경 최적화 등의 과제가 남았다고 생각한다. 또한 사용자를 추구한 나머지 관리자에 대한 부분은 현재 거의 개발되어 있지 않은게 아쉬운 부분이다.

향후 개발에서는 생성 **문제의 품질**을 더욱 높이고, **듣기 영역** 생성 문제를 추가하여 TOPIK 시험 전반을 아우르는 완성도를 높일 계획이다. 이를 통해 지속적으로 증가하는 유학생 공급과 한국어 교육에 필수적인 플랫폼이 될 것이라 생각한다.

6. 참고 문헌

- [1] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei, "Language Models are Few-Shot Learners," *Advances in Neural Information Processing Systems*, Vol. 33, pp. 1877-1901, 2020.
- [2] K. Warstadt, A. Parrish, H. Liu, A. Mohananey, W. Peng, S. F. Wang, and S. R. Bowman, "Learning Which Features Matter: RoBERTa Acquires a Preference for Linguistic Generalizations (Eventually)," *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 217-235, 2020.
- [3] S. Guo, X. Wang, and Y. Ban, "Automatic Generation of Medical Imaging Reports Using a Hierarchical Attention Network," *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*, pp. 3322-3332, 2021.
- [4] Y. Wang, X. Liu, and S. Shi, "Deep Learning for Automatic Mathematical Problem Generation," *Proceedings of the 2020 IEEE International Conference on Big Data (Big Data)*, pp. 5459-5468, 2020.
- [5] D. Kumar, A. Jain, and R. Kumari, "Automatic Question Generation from Text: A Comprehensive Survey," *International Journal of Information Technology*, Vol. 13, No. 4, pp. 1533-1546, 2021.