

2024 AI 아이디어 경진대회(전 분야)

제 목	단국대학교 행정업무 보조 챗봇: DANI-L (DAN AI LLM)			
팀 명	레볼루션		구성인원	총 3명 (팀 대표 포함)
구 분	성 명	대학 / 학과	학번 / 학년	학교메일
대표학생	정다훈	소프트웨어융합대학 /소프트웨어학과	32204041 3학년	32204041@dankook.ac.kr
참여학생	최예림	소프트웨어융합대학 /소프트웨어학과	32224684 3학년	choiyerim@dankook.ac.kr
	민유진	소프트웨어융합대학 /정보통계학과	32221598 3학년	minyujin951225@dankook.ac.kr
대회 참가 동의				
<p>최종 선발 시 2024년 AI 아이디어 경진대회(전 분야)에 참가하며 대회에서 요구하는 일정에 맞추어 활동에 성실하게 참여하겠습니다.</p> <p><input checked="" type="checkbox"/> 동의함 <input type="checkbox"/> 동의하지 않음 (※동의하지 않을 시, 경진대회 참가 불가)</p>				
아이디어 독창성 확인				
<p>본 아이디어는 타 대회에서 수상하였거나 외부에 공개된 것이 아니고 본 대회에만 제출되는 것임을 확인합니다.</p> <p><input checked="" type="checkbox"/> 예 <input type="checkbox"/> 아니오 (※ ‘예’가 아닌 경우, 경진대회 참가 불가)</p>				
<p>2024 AI 아이디어 경진대회(전 분야) 참가를 신청합니다.</p> <p style="text-align: right;">2024 . 11 . 24 .</p> <p>[붙임1] 아이디어 설명서(필수) [붙임2] 개인정보 수집 · 이용 · 제공 동의서(필수)</p> <p style="text-align: right;">대표학생 : 정다훈 </p>				

【개인정보 수집·이용·제공 동의 안내】

「개인정보보호법」제15조제1항제1호, 제17조제1항제1호, 제24조제1항제1호에 따라 아래와 같이 개인정보의 수집·이용·제공에 관하여 귀하의 동의를 얻고자 합니다.

□ 개인정보 수집·이용 내역

항목	수집·이용 목적	보유기간
소속 및 성명, 학과, 학번, 학년, 이메일	<ul style="list-style-type: none">· 「2024 AI 아이디어 경진대회(전 분야)」 대회 프로그램 운영·관리 목적· 경진대회의 홍보, 성과 확산 등을 위한 부분 혹은 전체 과정의 영상 촬영· 연구 저작권을 침해하지 않는 온라인 플랫폼 비영리적 목적의 노출	접수마감일로부터 6개월 이내




※ 정보주체는 「2024 AI 아이디어 경진대회(전 분야)」에 개인정보 수집·이용·제공의 동의를 거부할 권리가 있습니다.

그러나, 개인정보 수집·이용·제공에 동의하지 않을 경우, 본 프로그램의 선발대상에서 제외될 수 있습니다.

☞ 본인은 「2024 AI 아이디어 경진대회(전 분야)」에서 본인의 개인정보를 수집·이용·제공하는 것에 동의합니다.

동의	<input checked="" type="checkbox"/>	미동의	<input type="checkbox"/>
----	-------------------------------------	-----	--------------------------

2024년 11 월 24 일

본인(참가팀 대표)	정다훈	
본인(팀원)	최예림	
본인(팀원)	민유진	

[붙임1] 아이디어 설명서(필수)

[작성 요령]

- 서술형으로 작성하며, 11pt 줄간격 150% 기준 A4 10쪽 이내
- 기술, 자원, 시간 등을 고려하여 실제 구현이 가능한 아이디어를 제출
- 자유 양식이며 아래 목차는 반드시 포함

[작성 목차]

1. 주제명
2. 제안 배경
3. 내용 요약
4. 제안 내용 및 기존 사례 대비 장점(독창성)
5. 실현 가능성 및 활용성(구체적으로 설명)
6. 기대효과 및 사회적 가치

2024 AI 아이디어 경진대회 (전분야)



제목	단국대학교 행정업무 보조 챗봇: 다니엘(DANI-L, Dankook AI LLM)
----	---

팀 명: 레볼루션
대표학생: 정다훈
참여학생: 민유진, 최예림
제 출 일: 2024.11.24

목차

1. 주제명	1
2. 제안 배경	1
3. 내용 요약	1
4. 제안 내용 및 사례대비 장점	1
(1) 학생 - 질의응답 가능	1
(2) 교직원 - 업무 보조 가능	2
(3) 기존 사례비교: 단아이 DanAI	3
5. 실현 가능성 및 활용성	4
(1) 챗봇 구현과 관련된 개념 정리	4
(2) 챗봇 구현 계획	5
(3) 활용성	7
6. 기대효과 및 사회적 가치	7

1. 주제명

단국대학교 행정업무 보조 챗봇: 다니엘(DANI-L, Dankook AI LLM)

2. 제안 배경

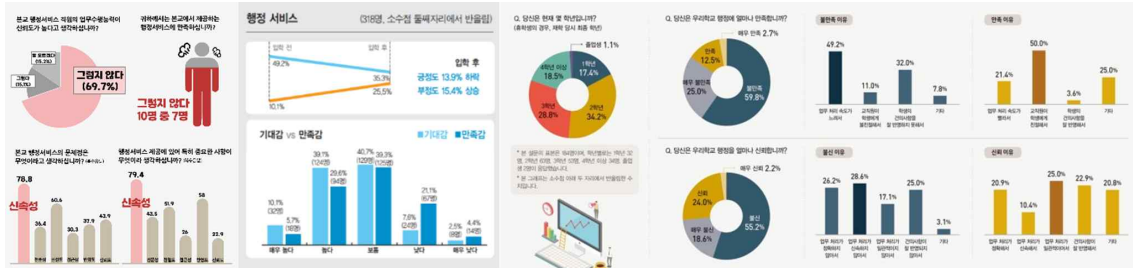


그림 1. 왼쪽부터 인하대학신문, 중대신문, 한성대신문의 행정 서비스 만족도 통계

각 대학에서는 소속 학생들을 대상으로 행정 서비스 만족도 조사를 실시하고 있다. 그러나 인하대학교, 중앙대학교, 한성대학교를 비롯한 많은 대학에서 학생들이 행정 서비스에 불만을 느끼는 경우가 많다. 주된 이유는 민원 처리가 오래 걸리거나, 담당 업무가 아니면 도움을 받을 수 없는 상황이 자주 발생하기 때문이다. 단국대학교 역시 이러한 문제에서 자유롭지 않다. 학과 사무실의 인원 교체가 잦아 업무 숙지가 부족한 경우가 많아, 학생들이 학과 행정 관련 질문에 대해 정확하고 신속한 답변을 받지 못하는 일이 빈번하다. 특히 담당자가 자리를 비운 경우 해당 행정 업무는 담당자가 복귀할 때까지 처리가 지연된다. 이러한 문제를 해결하기 위해 행정 업무를 보조할 수 있는 챗봇 도입을 제안한다.

3. 내용 요약

학생들은 행정 관련 정보를 얻기 위해 단국대학교 홈페이지와 학과 홈페이지 등 여러 웹페이지를 방문해야 한다. 그러나 원하는 정보를 찾지 못할 경우 결국 학과 사무실에 전화를 걸어 문의해야 하는 상황이 발생한다. 예를 들어, 학번에 따라 졸업 요건이 달라지는 것처럼 학생의 학사 정보에 따라 답변이 달라질 수 있어 전화 상담 중에도 실시간으로 정확한 정보를 제공하기 어려운 경우가 많다.

또한, 제출 서류, 행정 업무 방식, 학교 교칙 등이 변경될 때 최신 정보를 파악하는 데 어려움이 따른다. 더불어 학과 사무실은 담당 업무와 담당자가 자주 변경되는 특성으로 인해 행정 업무의 정확성과 효율성에 부정적인 영향을 미칠 수 있다. 이로 인해 업무 과중이 발생하며, 학생들에게 원활한 행정 서비스를 제공하는 데 한계가 생긴다. 이러한 문제를 해결하고 학과 사무실 행정 시스템의 부족한 점을 보완하기 위해 RAG를 활용한 행정 업무 보조 챗봇 도입을 제안한다.

4. 제안 내용 및 사례대비 장점

(1) 학생 - 질의응답 가능

챗봇은 행정, 교칙, 개설 과목, 공지사항 등 학교 행정과 관련된 데이터를 바탕으로 질의응답 서비스를 제공하며, 학생들이 필요한 정보를 빠르게 확인하고 문제를 해결할 수 있도록 돕는다. 데이터베이스에 저장된 행정 정보를 활용하여 사용자의 질문에 적합한 답변을 제공하며, 필요한 정보가 데이터베이스에 없는 경우에는 관련 부서의 연락처를 안내해 추가적인 지원을 받을 수 있도록 한다.

또한, 데이터베이스는 정기적으로 최신 정보를 반영하도록 업데이트되어, 변경된 학교 행정 사항이나 새로운 공지사항 등을 실시간으로 제공한다. 이를 통해 학생들에게 신뢰도 높은 최신 정보를 제공하며, 사용자 편의를 증대시킨다.

학습 데이터

학생 질의응답을 위한 학습 데이터로는 학교 행정과 관련된 답변을 위해 단국대학교 학칙, 공지사항, VOC(FAQ) 답변 내용을 활용할 수 있다. 또한, 학교 수업과 관련된 질문에 답변하기 위해 학과별 개설 과목 정보, 졸업 요건, 학사일정 등의 데이터를 포함할 수 있다.

만약 질문에 대한 정보가 데이터베이스에 없는 경우, 해당 업무를 담당하는 부서 또는 담당자의 연락처를 제공한다. 이를 위해 각 업무를 담당하는 부서와 담당 직원의 연락처를 데이터베이스에 체계적으로 저장한다.

질문 예시

학생들이 챗봇을 통해 할 수 있는 질문의 예시는 다음과 같다.

- **과목 정보 질문:** A 과목은 전공 선택 과목인가요, 아니면 POSE 과목인가요?
- **학사 일정 질문:** 2024년 2학기 성적 장학금 신청을 위한 공인영어성적 제출 마감일은 언제인가요?
- **졸업 관련 질문:** 현재 졸업 요건 중 부족한 점이 무엇인가요?
- **공지사항 질문:** 2025년 1학기 교환학생 신청 양식 파일을 제공해 주시겠어요?

(2) 교직원 - 업무 보조 가능

학생들이 담당 직원과 전화 상담을 통해 문제를 해결하려는 경우, 챗봇은 상담 중인 학생의 정보와 관련된 적절한 데이터를 검색하여 실시간으로 제공할 수 있다. 또한, 챗봇은 업무 매뉴얼을 학습하여 담당자가 익숙하지 않은 업무에 대한 가이드를 제공함으로써 원활하고 신속한 인수인계를 지원한다.

추가적으로, 학생이 행정 업무와 관련된 서류를 제출할 때 챗봇은 서류의 형식과 누락된 내용 여부를 검토하여 담당자에게 전달한다.

질문 예시

- **전화 상담 지원**
 - "소프트웨어학과 20학번의 졸업 요건은 어떻게 되나요?"
 - "교환학생 프로그램에 대한 신청 자격과 일정은 어떻게 되나요?"
- **업무 가이드 제공(신속한 인수인계)**
 - "이 서류를 처리하려면 어떤 절차를 따라야 하나요?"

- "학사 일정 관련 공문 작성은 어떻게 하면 되나요?"

(3) 기존 사례비교: 단아이 DanAI

단아이

현재 단국대학교는 단국인(학생 및 교수자)을 위한 인공지능 기반 맞춤형 교육 지원 서비스인 교육지원비서, 즉 단아이와 같은 챗봇 서비스를 제공하고 있다. 그러나 학생들의 질문에 대해 챗봇이 적절한 답변을 제공하지 못하는 것을 관찰할 수 있다.



그림 2. 단아이 (Dan.i) 실행 화면

단아이 서비스에 "졸업까지 남은 학점"에 대해 질문했을 때 졸업과 관련된 학사 일정 정보를 제공하거나, "공인 영어 성적 신청 마감일"에 대한 질문에 졸업 인증 기준을 제공하는 등의 문제를 확인할 수 있다. 이와 같이 단아이는 몇 가지 문제를 가지고 있다.

단아이의 문제점

위의 내용을 기반으로 단아이의 문제점은 다음과 같이 정리할 수 있다.

1. 질문의 의도를 정확히 파악하지 못하고, 키워드만 추출하여 답변을 제공
2. 답변이 링크 형태로 제공되어, 챗봇으로서의 본연 기능을 충분히 수행하지 못함
3. 개인 맞춤형 챗봇이라는 소개와 달리, 사용자에게 적절한 답변을 제공하지 못함

단아이와의 차별점

본 보고서에서 제안하는 챗봇은 단아이의 단점을 보완하고, 학교의 최신 정보를 반영하여 정책 변경 시에도 신속하게 대응할 수 있도록 설계되었다. 또한, 학생 개인 데이터를 연동하여 맞춤형 정보를 제공하며, 단순히 링크 형식의 답변 대신 자연어 기반의 대화형 답변을 통해 보다 효과적이고 자연스러운 상담이 이루어지도록 구현할 예정이다.

- **정확한 답변 제공:** 질문의 맥락을 이해하지 못하고 키워드만 추출해 부정확한 답변을 제공하는 기존 단아이의 한계를 보완

- **자연어 기반 대화 지원:** 단순히 링크를 제공하는 형식에서 벗어나, 자연어를 활용한 대화형 답변을 제공하여 사용자 경험을 개선
- **최신 정보 반영:** 학교의 최신 정보를 신속히 반영하여, 정책이나 교칙 변경 시에도 즉각적으로 대응할 수 있도록 설계
- **개인 맞춤형 정보 제공:** 학생 개인 데이터를 연동하여 시간표, 학점 현황 등 맞춤형 정보를 제공함으로써 사용자에게 적합한 답변을 제공
- **효율적인 상담 구현:** 자연스럽고 효과적인 상담을 통해 단순한 정보 제공 이상의 실질적인 도움을 제공하도록 구현할 예정

5. 실현 가능성 및 활용성

(1) 챗봇 구현과 관련된 개념 정리

LLM

대규모 언어 모델(Large Language Model, LLM)은 인간처럼 텍스트를 이해하고 생성하도록 설계된 신경망이다. 방대한 텍스트 데이터와 수십억 개의 파라미터를 기반으로 훈련되며, 텍스트의 문맥과 구조를 이해하고 언어의 순차적 특성을 활용해 다음 단어를 예측하는 방식으로 작동한다. 이러한 기능을 통해 LLM은 논리적이고 의미 있는 텍스트 생성이 가능해진다.

RAG

LLM의 한계를 보완하기 위해 '지식 검색'과 '언어 생성'을 결합한 기술로 LLM이 최신 정보나 특정 도메인 지식을 제공하는 데 어려움을 해결하고자, 외부 지식 베이스에서 관련 정보를 검색하여 이를 기반으로 사실적이고 신뢰성 있는 답변을 생성한다. 또한, 검색된 정보의 출처를 명시함으로써 설명 가능성과 정확성을 높이고, 질의에 대한 맥락을 이해하여 추론 능력을 강화한다. RAG(Retrieval Augmented Generation)는 이러한 방식을 통해 LLM의 생성 능력과 사실 관계 파악 능력을 향상시키는 프레임워크를 의미한다.

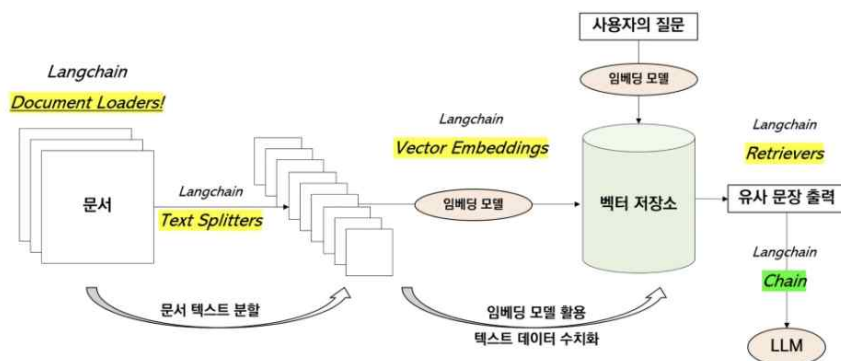


그림 3. RAG 작동 원리

RAG는 검색 단계와 생성 단계로 구분된다. 사용자가 입력한 쿼리를 수신하면, 검색 모델이 데이터베이스나 외부 소스에서 관련 정보를 검색하여 이를 고차원 벡터로

변환하고 벡터 데이터베이스에 저장한다. 이후 입력 쿼리와 관련성을 기준으로 문서 순위를 매기고 가장 높은 점수를 받은 문서를 선택한다. 그 후, LLM은 검색된 정보를 바탕으로 텍스트 응답을 생성하여 답변을 제공한다. 이 과정을 통해 RAG는 모델의 응답 정확도를 높이고, 외부 지식을 활용해 실용적인 답변을 생성할 수 있도록 돕는다.

(2) 챗봇 구현 계획

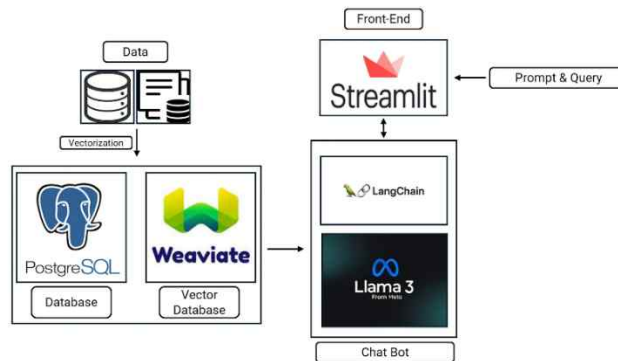


그림 4. 챗봇 아키텍처

데이터베이스: PostgreSQL

데이터 읽기 및 쓰기 작업을 지원하는 데이터베이스로는 PostgreSQL을 사용한다. PostgreSQL은 고성능 읽기 작업을 지원하며, 캐싱 및 검색 최적화가 잘되어 있어 수천 명 이상의 동시 접속자에게도 원활한 서비스를 제공한다. 또한, 확장성이 뛰어나 추후 챗봇에 새로운 기능을 추가할 때에도 유리하다.

벡터 데이터베이스: Weaviate

데이터베이스에 저장되는 데이터들을 벡터 임베딩하여 벡터로 저장하는 벡터 데이터베이스 프로그램으로는 Weaviate를 사용한다. 벡터 데이터베이스란 텍스트, 이미지, 비디오 등 비정형 데이터를 벡터 형태로 저장하고, 벡터 간 유사도를 기반으로 검색을 수행하는 데이터베이스를 말한다. Weaviate는 유사도 검색에 최적화되어 있고, 빠른 검색 속도를 지원하며 다양한 Machine Learning 모델과 자연스러운 통합을 지원한다. 이는 대규모 비정형 데이터를 다루면서 빠르고 정확한 검색 결과를 제공해야 하는 챗봇 애플리케이션에 적합하다.

LLM(Large Language Model): Llama 3

LLM로는 Llama 3를 사용한다. Llama 3는 Meta에서 개발한 최신 대규모 언어 모델로, 텍스트 데이터를 기반으로 자연어의 문맥과 의미를 학습하여 다양한 자연어 처리(NLP) 작업을 수행할 수 있다. 이 모델은 특히 학습 파라미터와 데이터 처리 속도 측면에서 최적화되어 있다. Llama 3는 텍스트 생성, 질문 응답, 요약 등 다양한 작업에 강점을 보이며, 높은 정확도로 유의미한 결과를 생성한다. 이러한 특성은 대규모 학사 데이터나 비정형 데이터를 다루는 챗봇 애플리케이션에서 자연스러운 대화와 정확한 정보 제공을 가능하게 하며, 사용자 경험을 크게 향상시킬 수 있다.

UI(User Interface): Streamlit

UI로는 Streamlit을 사용한다. Streamlit은 Python 기반의 웹 애플리케이션 프레임워크로, 데이터 시각화와 대화형 사용자 인터페이스를 간단하게 구축할 수 있는 도구이다. 코드를 간결하게 작성하면서도 강력한 기능을 제공하며, 데이터 분석, 모델 결과 시각화, 사용자 입력 처리 등 다양한 작업에 적합하다. Streamlit은 실시간으로 데이터를 업데이트하고 대화형 요소(슬라이더, 버튼, 입력 필드 등)를 제공하여 사용자와 애플리케이션 간의 상호작용을 원활하게 한다. 이를 통해 챗봇 애플리케이션의 결과를 직관적으로 확인하고, 사용자 경험을 개선하며, 손쉽게 커스터마이징된 UI를 구축할 수 있다. 이러한 특성 덕분에 Streamlit은 챗봇의 프론트엔드 개발에 효율적이고 실용적인 선택이다.

학생 Chatbot 동작 순서

1. 질문 접수 및 초기 처리

- 학생들이 입력한 질문과 학사 정보를 수집한다.
- 질문을 담당 부서별로 분류하거나, 적합한 데이터베이스 및 부서를 식별한다.

2. 질문과 데이터 비교

- 질문을 임베딩(Embedding) 벡터로 변환하고, 벡터 스토어에 저장된 데이터와 유사도 검사를 수행한다.
- 임베딩 벡터 스토어에 적합한 데이터가 존재하지 않을 경우:
 - 질문을 파인튜닝된 모델(Llama 3 또는 유사한 LLM)로 처리하여 담당 부서를 추론한다.
 - 부서 정보 데이터베이스에서 해당 부서의 담당자 및 연락처를 검색하여 제공한다.

3. 데이터 기반 답변 생성

- 유사도 검사에서 반환된 데이터를 기반으로 Retriever-augmented Generation (RAG)의 Retriever를 통해 적절한 문서를 검색한다.
- 검색된 데이터를 Llama 3(LLM)에 입력하여 답변을 생성한다.
- 답변 생성 시, 정보가 참조된 문서의 출처(문서 이름, 내용 요약 등)를 함께 제공한다.

4. 답변 전달

- 학생에게 생성된 답변을 제공한다.
- 필요 시 부가적인 문서나 연락처 등의 추가 정보를 함께 제공한다.

교직원 Chatbot 동작 순서

1. 파일 업로드 및 초기 처리

- 학생이 파일을 업로드하면, 챗봇이 해당 파일을 수신한다.
- LLM(대규모 언어 모델)을 사용하여 파일이 서류 평가 기준에 따라 적절하게 작성되었는지 평가한다.
 - 평가 기준: 별도로 설계된 서류 평가 기준 데이터를 활용한다.

RAG 기반의 챗봇 시스템은 행정 업무의 자동화를 지원하여 교직원들이 학생 상담, 서류 처리, 업무 인수인계에 소요되는 시간을 크게 절감할 수 있다. 이를 통해 행정 효율성이 향상되고, 학교 행정의 전반적인 속도와 정확성이 개선되게 된다.

학생 편의성 증대

학생들은 학교 행정과 관련된 다양한 질문을 실시간으로 해결할 수 있어, 불필요한 대기 시간을 줄이고 필요한 정보를 신속하게 얻을 수 있게 된다. 또한, 학생 개인 정보를 연동하여 보다 맞춤형 답변을 제공함으로써 개인화된 서비스를 경험할 수 있다.

정확한 정보 제공

RAG를 활용한 기술은 학교의 최신 정보를 실시간으로 반영하고, 정책 변경 시 신속하게 업데이트할 수 있어 학생들에게 정확한 정보를 제공해준다. 기존의 단아이 시스템에서 발생했던 부정확한 답변 문제를 해결하고, 데이터베이스의 정보를 결합하여 더 신뢰할 수 있는 답변을 제공할 수 있다.

업무 인수인계 및 교육 지원

교직원들은 업무 메뉴얼을 학습한 LLM을 통해 각종 업무 절차를 자동으로 안내받고, 적절한 가이드를 제공받음으로써 행정 업무를 빠르게 숙지할 수 있다. 이를 통해 새로운 직원들도 원활하게 업무를 습득할 수 있으며, 효율적인 업무 진행과 빠른 인수인계가 가능해진다. 결과적으로 교직원들의 업무 숙련도를 높이는 데 도움이 된다.

학생 - 교직원 간 소통 증진

학생과 교직원 간의 소통이 더욱 원활해지고, 상호 작용의 질이 향상된다. 학생은 챗봇을 통해 실시간으로 질문을 하고 답변을 받으며, 교직원은 챗봇을 통해 학생의 요구사항을 더욱 신속하게 처리할 수 있게 된다.

결론

결과적으로, 학생들에게는 편리한 방식으로 더 정확한 정보를 제공하고, 교직원들에게는 업무 숙련도를 향상시킬 기회를 제공한다. 이를 통해 단국대학교 서비스의 질을 향상시키는 사회적 가치를 창출할 것으로 전망된다.