

KT AICC 솔루션

NLP (Natural Language Processing) 소개자료



클라우드편

DIGICO KT

Contents

1 KT NLP 솔루션 소개

2 NLP 솔루션 구성

3 NLP 솔루션 기능

4 NLP 특징점

5 NLP 활용방안

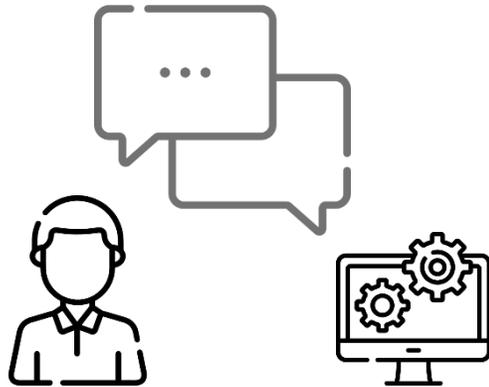
별첨

01 KT NLP 솔루션 소개

KT NLP 솔루션은 딥러닝 및 기계학습 알고리즘을 활용하여 의도 분류 및 개체명 인식 기능을 제공합니다.

✓ NLP(Natural Language Processing, 자연어처리기술) 란?

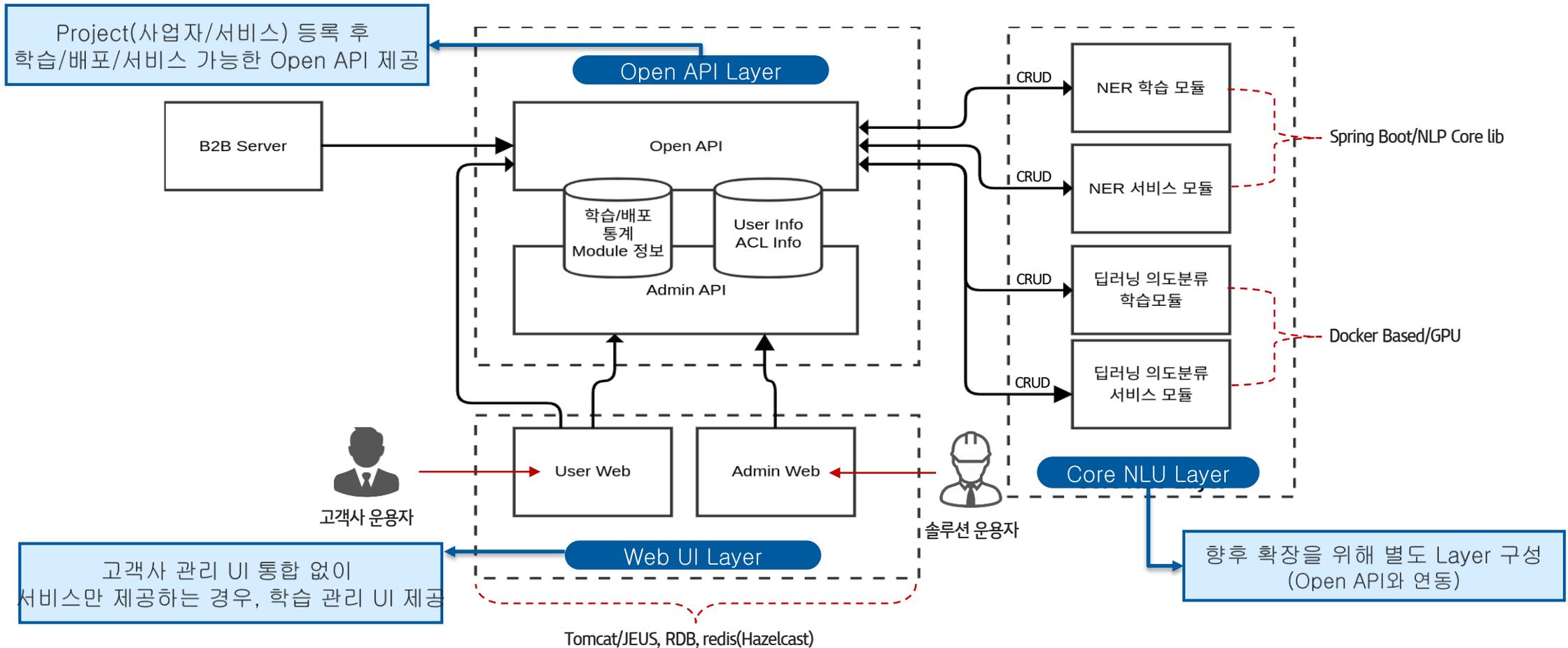
:컴퓨터와 인간언어 간 상호 작용을 하기 위한 기술로, 자연어에서 의미 있는 정보를 추출, 이해, 분석하여 컴퓨터가 처리 할 수 있도록 하는 기술



자연어 전처리	전	후
띄어쓰기 처리	개설하려면어떻게	개설하려면 어떻게
형식파괴 수정	하나용	하나요
동의어 변환	K	KT
오타 정정	개좌	계좌
복합명사처리	KT 은행	KT은행
불용어 제거	ㅎㅎ	삭제

02 NLP 솔루션 구성

KT NLP 솔루션은 컴포넌트 단위의 구성을 하여 성능을 보장하고 사용자의 편의성을 증대시킬 수 있도록 개발하였습니다.
(Open API, UI, Core NLU Layer로 구성)



03 NLP 솔루션 기능

자체 개발한 자연어 전처리 엔진과 약 80만개의 형태소 사전을 포함한 다수의 사전 데이터를 이용하여 여러가지 기능들을 제공합니다.

1 사전 관리 기능

- 형태소 분석 결과 확인 기능 제공
- 복합명사, 고유명사, 축약어 등 전문 용어 사전 관리
- 다양한 상품 등록 및 관리 필요 (용어 차이 고려)
- Stopword (불용어, 비속어) 관리
- 유의어 사전 관리
- 동일한 의도 분석되어야 하는 키워드 유의어 등록 및 관리



2 개체명(Entity) 인식 기능

- 기계학습 알고리즘(CRF)을 활용한 KT 자체 형태소 분석/사전 제공
- 개체명관리를 통한 정확한 사용자 의도 파악 지원
- 개체명 인식 검증 테스트 기능 제공
- 딥러닝 및 Rule based 모델 적용
- 개체명에 대한 유의어 구축을 통해 인식도 증대

3 의도분류(Intent) 기능

- 대용량 언어모델(KT-RoBERTa) 기반 신경망 학습
- 사용자 발화 의도 파악하여, 수행할 기능 판별
- 자연어처리 기술을 통해 의도 분류 관리
- 개체명 태깅, 형태소 분석, 토큰 분리, 유의어 등
- CPU 기반의 딥러닝 Paraphrasing 기술 구동

4 학습문장관리(언어자원처리) 기능

- 구축한 데이터 및 사전, 로그 기록 관리 지원
- 지속적인 성능 향상 관리
- 시나리오 분석 설계 자료 활용
- 로그 기록으로 자동완성 지원하여 편의성 지원

5 부가 Annotator 관리 기능

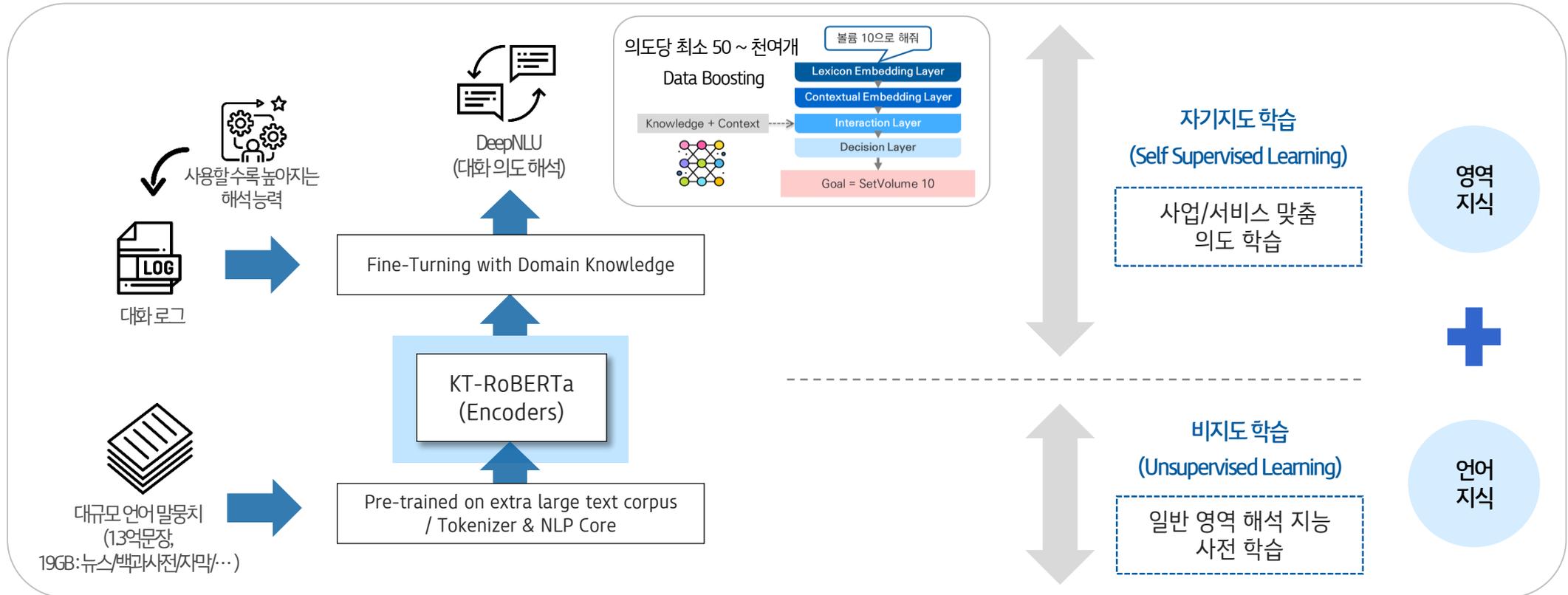
- 한글/숫자 금액 분석 처리 지원
- 숫자 인식 보정 처리
- 부가정보 태깅 지원
- 인입되는 질의 및 정보에 대해 분석자료 활용
- 개체명 등록 지원
- 특정 태깅 정보 선택 시 개체명 사전 등록

04 NLP 특징점

(1) DeepNLU(Natural Language Understanding) 의도 해석이 가능합니다.

* DeepNLU(Natural Language Understanding) : 대규모 언어데이터에서 학습한 신경망 기반의 해석기

- ✓ 대규모 신경망에 기반하여 확률적으로 의도 해석 가능하고 발화 실수나 인식 오류에 대해 강건
- ✓ 사전/N-gram패턴/규칙에 기반한 전/후처리를 Hybrid 조합하여 성능 최적화 가능



05 NLP 활용 방안

KT NLP 솔루션은 'AI 컨텍센터'와 'AI 비서'에 활용 가능합니다.

AI 컨텍센터
(챗봇, 콜봇, 상담 어시스트 등)



챗봇

다양한 채널을 통해 입력된 고객의 요청에
최적의 답변을 대화형으로 제공



콜봇

전화상담(문의/예약 등) 고객을 대상으로
가상상담원이 자동으로 상담대응

AI 비서
(인공지능 스피커, 예약봇 등)

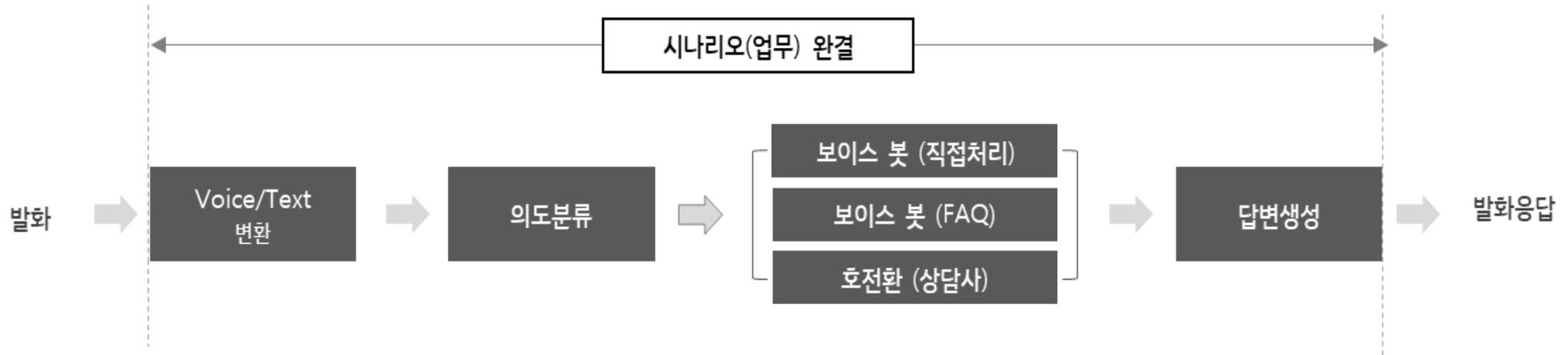


음성처리 솔루션 적용으로 음성인식 성능 향상
호출어 응답소리 변경
맞춤형 서비스 제안
- 고객의 상황을 먼저 캐치하여,
기가지니가 먼저 고객에게 알맞은 서비스를 제안

솔루션 Layer	Bot Platform			AI Platform					
	Voice GW	SOE		STT	TTS	TA	MRC	NLP	
	SIP I/F	WEB View	API/Router Mgmt	실시간 STT 분석	오디오 인코딩	내부 데이터 수집 모듈	FAQ 엔진	Intent 학습	
	SOE I/F	External CH I/F	Log Mgmt	준실시간 STT 분석	인코더 모델		MRC 엔진	개체명(Entity) 학습	
	PBX I/F	NLP I/F Mgmt	Security	음원 수집/분배	디코더 모델	TA 엔진	Passage 서치 엔진	Dialog 개발	
	PDS I/F	Dialog Mgmt	Statistics Mgmt	음성 인식/변환	보코더 모델	강성 분석	테이블 QA 엔진	Context Variable (명령어)	
	녹취 I/F	Connection/Session	기타 모듈	언어모델 학습	음성합성 엔진	규칙 처리	ML 학습	대시보드	
				통합관리	엔진 및 채널 관리	의미 분석	DL 학습		
						형태소 분석			

별첨 NLP 성능

KT 콜센터 적용 NLP의 품질 기준은 E2E 완결을 기준으로 합니다.



$$\text{STT 성능} + \text{NLP 성능} + \text{대화/시나리오 성공률 인프라 성공률} = ?$$

DIGICO **KT**