

KT AICC TTS(Text to Speech) 솔루션



클라우드편

DIGICO KT

Contents

-
- 1 도입배경
 - 2 솔루션 개요
 - 3 솔루션 특징점
 - 4 솔루션 적용 기대효과
 - 5 별첨
-

01 도입 배경

언택트 트렌드에 맞는 콜센터 상담 환경 구축 및 전문상담가에 대한 고객의 니즈가 증가하고 있음에 따라, 기업의 서비스 경쟁력을 강화하고 업무 효율성을 확보할 수 있는 차별화된 솔루션이 필요합니다.

Needs

고객 편의성
극대화

코로나 19 장기화에 따른 비대면 업무처리 선호

차별화된 서비스 기대

AI 기반
상담시스템

빠르고 편리한 음성 안내 서비스 요구

자동화 시스템에 대한 니즈 확대

업무 자동화

Robotics로 금융 업무 자동화 요구

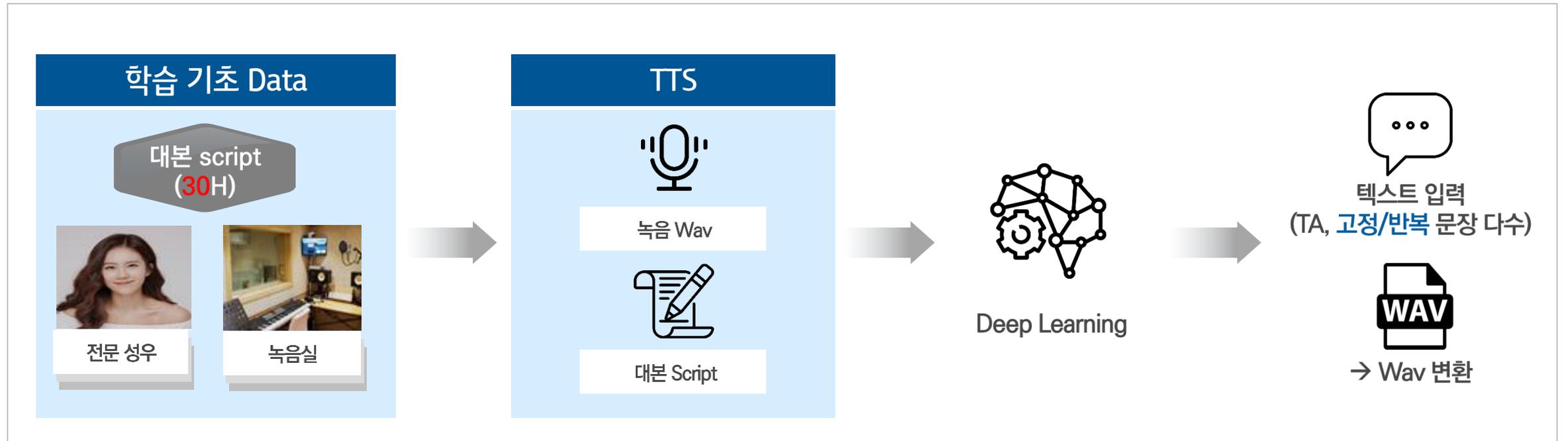
정형화된 대량 상담 업무 처리 필요

고객 니즈에 맞춰 AI 기반 고객센터에 응대 가능한 TTS(Text to Speech) 솔루션 필요

02 솔루션 개요

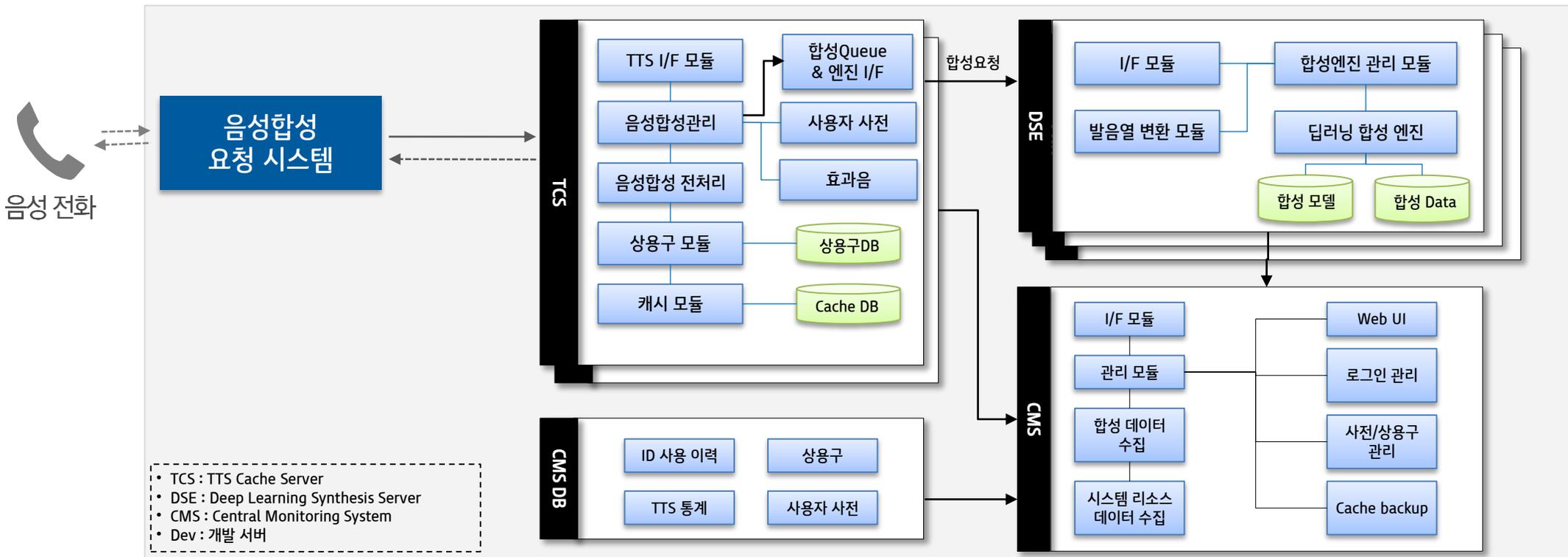
TTS 솔루션은 문자를 사람의 음성으로 변환하는 기술로 KT 자체 개발엔진으로 자연스러운 음성 합성 품질을 제공하는 시스템입니다.

AICC플랫폼



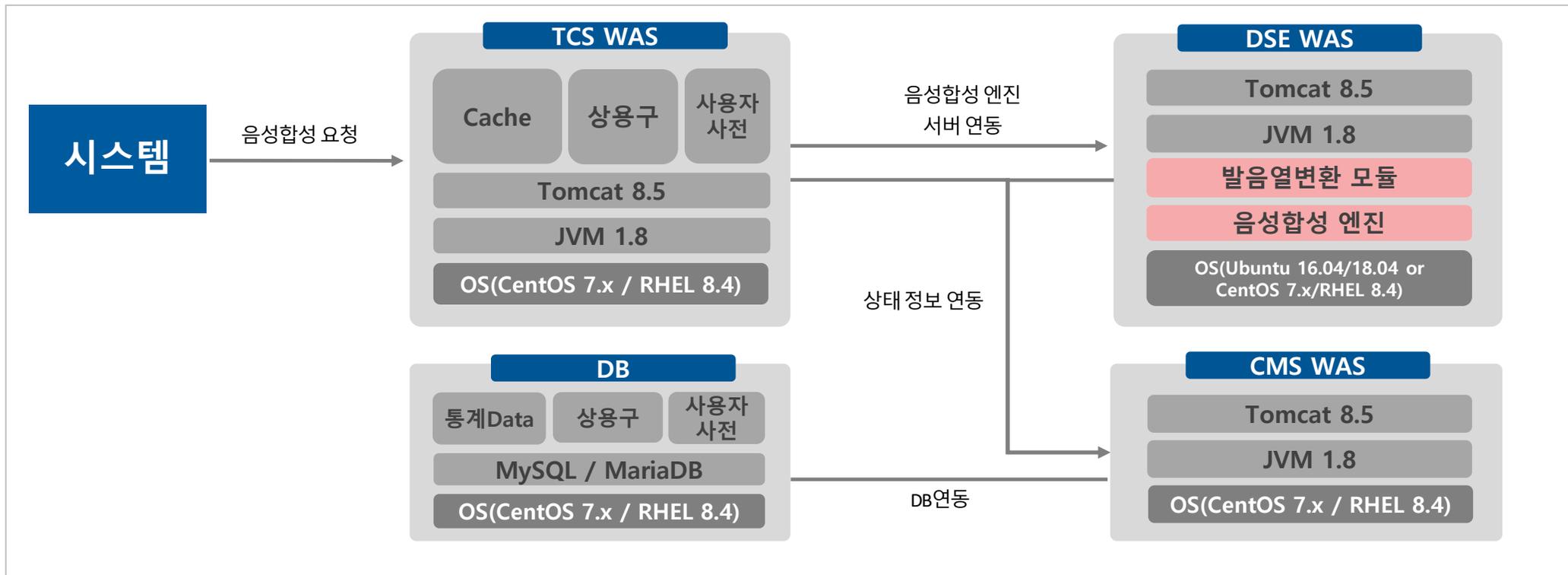
02 솔루션 개요 (TTS 서버 구성)

효과적인 음성합성성능 향상을 위한 상용구, 캐시의 활용 및 효과적인 음성합성엔진서버의 활용
메모리 기반 저장소의 활용을 통한 속도 및 I/O 사용률 극대화



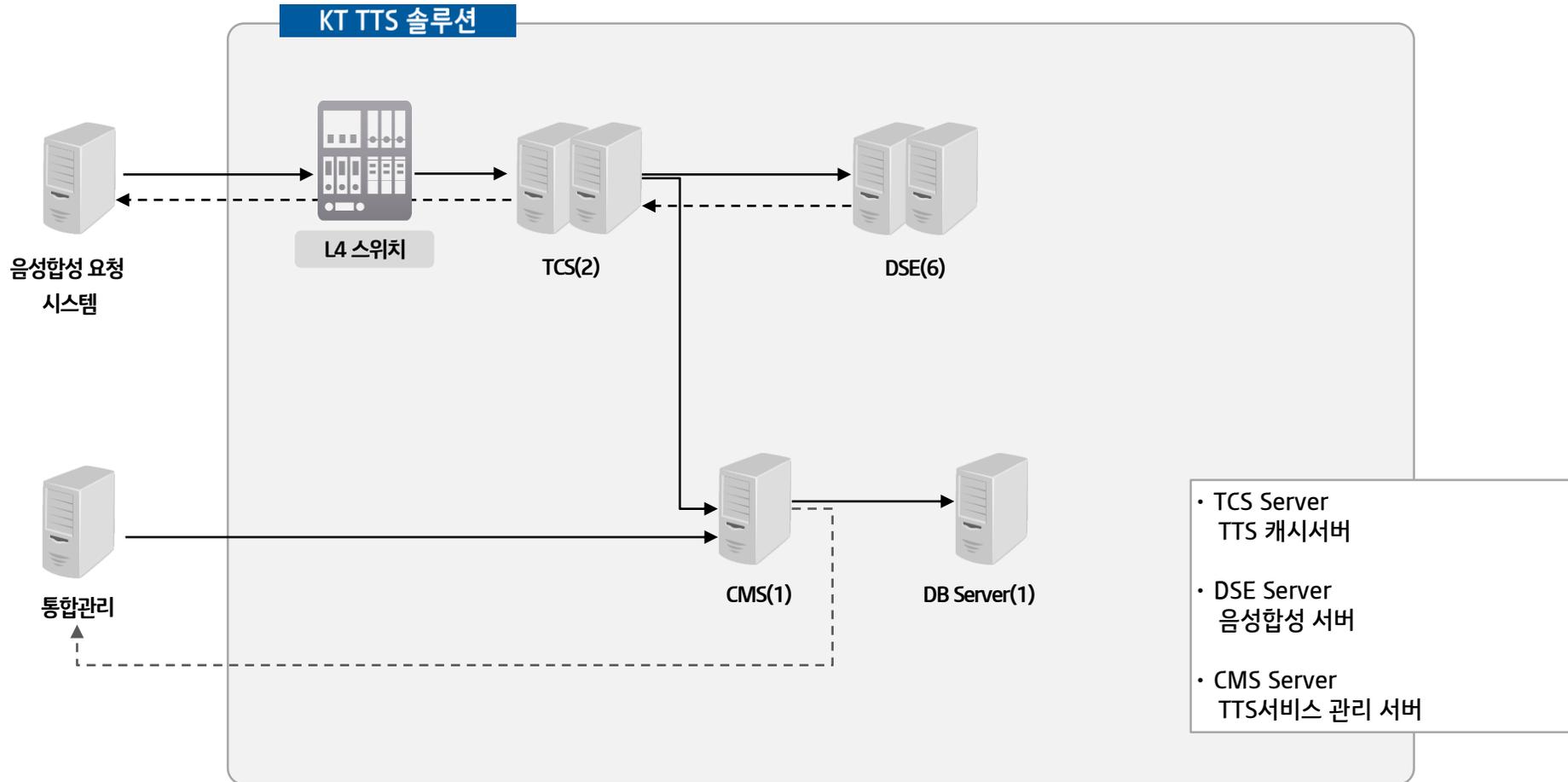
02 솔루션 개요 (TTS S/W 구성)

- ✓ TCS서버 : 음성합성 요청에 대한 전처리 기능과 캐시/상용구를 이용한 음성합성 및 음성합성엔진 서버에 대한 제어를 수행
- ✓ DSE서버 : 실제 음성합성을 수행
- ✓ CMS 서버 : 콘텐츠 관리 및 관제 담당



02 솔루션 개요 (처리 프로세스)

음성합성 요청 시스템(녹취 시스템, 보이스봇 GW 등) 으로부터 요청을 받아, L4 스위치에서 캐쉬 서버로 요청 분배하여 합성합니다.



03 솔루션 특징점

KT 자체 개발 엔진(CPU기반 DNN TTS)으로 음성합성 성능 향상을 위해 상용구 및 캐시를 활용하여 정교한 언어처리 및 자연스러운 합성 품질 제공합니다.

텍스트 데이터의 자연스러운 음성합성

- 문맥에 맞게 숫자, 날짜, 금액 등을 자연스럽게 읽어주는 음성합성음 제어 기술 적용
- 특수문자, 알파벳, 숫자를 자연스럽게 읽어주는 이메일 태그 기능
- 서비스별 문장 등록 및 음질 확인 가능



CPU기반 딥러닝 TTS 기술 적용 (GPU 미사용)

- GPU기반 TTS 대비 운영비용 절약
- DNN 기반 고품질 음성합성 제공
- 짧아진 신규 보이스 개발 시 필요한 녹음 Data 시간: 30 시간

고객별 사용자 정의/합성 규칙 생성 및 신규 단어/불완전한 발화 튜닝

- 사용자 사전을 통한 신규단어/불완전 발화 튜닝
예) 코로나19 → 코로나일구, 63빌딩 → 육삼빌딩
- 사용자 사전을 통해 단어/문장 독음 설정 관리

상용구 및 캐시 기능을 통한 음성합성 성능 향상

- 인트로 멘트 등 반복사용 문장은 상용구 설정하고 캐시에 저장
- 상용구는 TTS 변환 시 생성과정 생략
→ 시간 단축시스템 리소스 절약

03 솔루션 특징점

TTS 솔루션은 KT가 직접개발한 딥러닝 기반 솔루션으로 금융 서비스에 최적화된 고품질 음성 제공합니다.
(고객 조사 기반 최적화된 합성음 채택 등 서비스에 가장 적합한 목소리 제공)

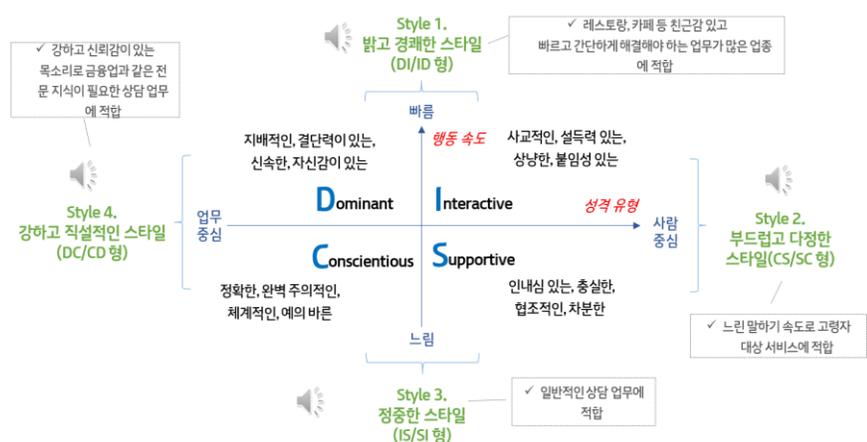
다수의 레퍼런스를 보유한 검증된 P-TTS

- 
 - 국내최초 셀럽 P-TTS 기반 AI스피커 서비스 출시
 - 실시간 랭킹 퀴즈 서비스 “박명수를 이겨라”
- 
 - 기가지니 키즈 서비스에 핑크퐁 캐릭터 TTS 음성발화
- 
 - “내목소리동화“ : 부모 목소리 300문장 녹음으로 다양한 동화책을 자녀에게 들려주는 서비스
- 
 - 적은 음성데이터로 개인화 음성합성 제공
 - 청각 장애인의 목소리 합성 재현
- 
 - 100번 공감 표현형 상담 보이스를 별도 제작
 - 공감 우수상담사 목소리로 제작

고객 맞춤형 TTS 보이스 제작 지원

- ##### 맞춤형 보이스 제작 프로세스

 - 1) 컨셉 정의: 목소리 페르소나 정의 및 스타일 선정
 - 2) 녹음 준비: 녹음 가이드 및 고객사 맞춤 스크립트 제작
 - 3) 녹음: 성우 녹음 및 발음열 변환
 - 4) 딥러닝 훈련: 딥러닝 훈련 및 추가 보완
- ##### User Research 기반 서비스 도메인 별 보이스 스타일 선정 가이드라인



이 매트릭스는 '행동 속도' (빠름 vs 느림)와 '성격 유형' (강령적 vs 사람 중심)를 축으로 보이스 스타일을 분류합니다.

 - Style 1. 밝고 경쾌한 스타일 (DI/ID 형)**: 레스토랑, 카페 등 친근감 있고 빠르고 간단하게 해결해야 하는 업무가 많은 업종에 적합.
 - Style 2. 부드럽고 다정한 스타일 (CS/SC 형)**: 느린 말하기 속도로 고객자 대상 서비스에 적합.
 - Style 3. 정중한 스타일 (IS/SI 형)**: 일반적인 상담 업무에 적합.
 - Style 4. 강하고 직설적인 스타일 (DC/CD 형)**: 강하고 신뢰감이 있는 목소리로 금융업과 같은 전문 지식이 필요한 상담 업무에 적합.

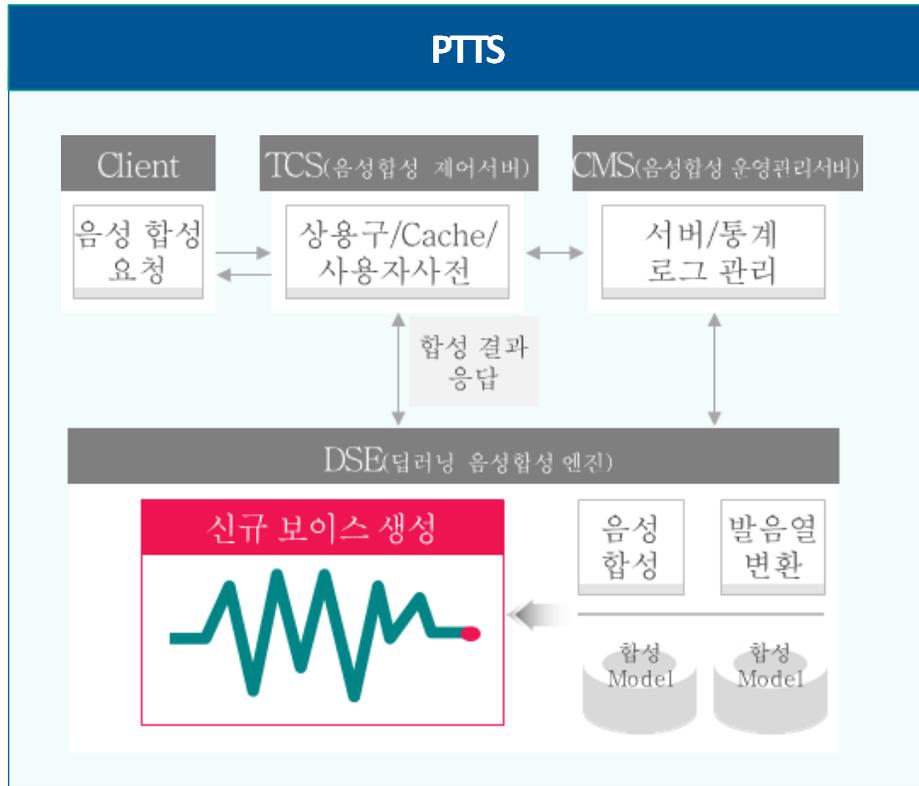
매트릭스 세부 설명:

 - 빠름 / 강령적 (Dominant):** 지배적인, 결단력이 있는, 신속한, 자신감이 있는.
 - 느림 / 강령적 (Conscientious):** 정확한, 완벽 주의적인, 체계적인, 예의 바른.
 - 빠름 / 사람 중심 (Interactive):** 사교적인, 설득력 있는, 상냥한, 불임성 있는.
 - 느림 / 사람 중심 (Supportive):** 인내심 있는, 충실한, 협조적인, 차분한.

04 솔루션 적용 기대 효과

TTS 솔루션은 딥러닝 기술을 이용하여 기업 고객 전용 보이스로 적용이 가능합니다.

- 아이덴티티와 일관성 있는 보이스 UX 제공함으로써 고객센터 응대의 품질 제고와 고객만족도 향상에 기여 합니다.



• 지원 보이스

- 여성2개(다솜, 세아), 남성2개(수호, 승완)

여성



다솜



세아

남성



수호



승완

• 고객사 전용 보이스 개발 지원

✓ 신규 보이스 개발

-> kt 기존 보이스가 아닌 별도 신규 보이스 제작 (신규 성우 선정 필요)

- 녹음 문장 수: 10,000~20,000 문장

- 녹음 소요 시간: 100~200시간

- 개발 비용: 약 1억원

- 개발 기간: 약 6~8개월

✓ 보이스 스타일 개발

-> kt 기존 보이스를 베이스로 스타일 변경 (ex: 기존 목소리에서 더 밝은 스타일)

- 녹음 문장 수: 3,000~3,500문장

- 녹음 소요 시간: 15~25시간

- 개발 비용: 약 2,000만원~3,000만원

- 개발 기간: 약 12주 (성우 녹음 및 딥러닝 학습 기간 포함)

별첨 KT TTS 보이스 스타일

기본 TTS 보이스

[여성 목소리]

다솜  - KT AI 콜센터 상담 사용 목소리
 - 일반적인 상담 업무에 적합
 - 친절하고 협조적인 느낌

세아  - 기가지니 사용 목소리
 - 다솜과 비교하여 억양 변화가 많고 음높이가 높은 목소리

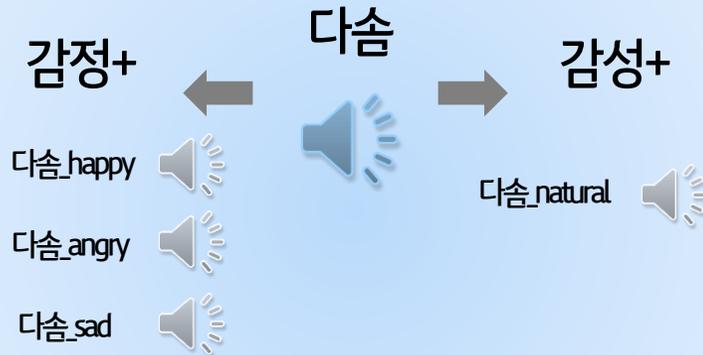
[남성 목소리]

수호  - 일반 남성에 비해 넓은 음역대
 - 약간 빠른 발화 속도로 적극적인 느낌

승완  - 수호와 비교하여 음높이가 낮음
 - 비교적 속도가 느려 차분하고 안정적인 느낌

보이스 스타일 개발

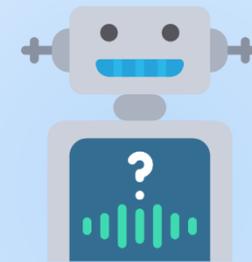
- 기존 KT 보이스를 베이스로 스타일 데이터 추가 학습하여 새로운 목소리 개발
- 효율적으로 새로운 목소리 TTS 개발



- 녹음 문장 수: 3,000 ~ 3,500문장
- 녹음 소요 시간: 15 ~ 20시간
- 개발 비용: 약 2천만원~3천만원
- 개발 기간: 약 12주

신규 보이스 개발

- 기존 KT 보이스가 아닌 별도 신규 보이스 제작
- 새로운 성우 선정 필요

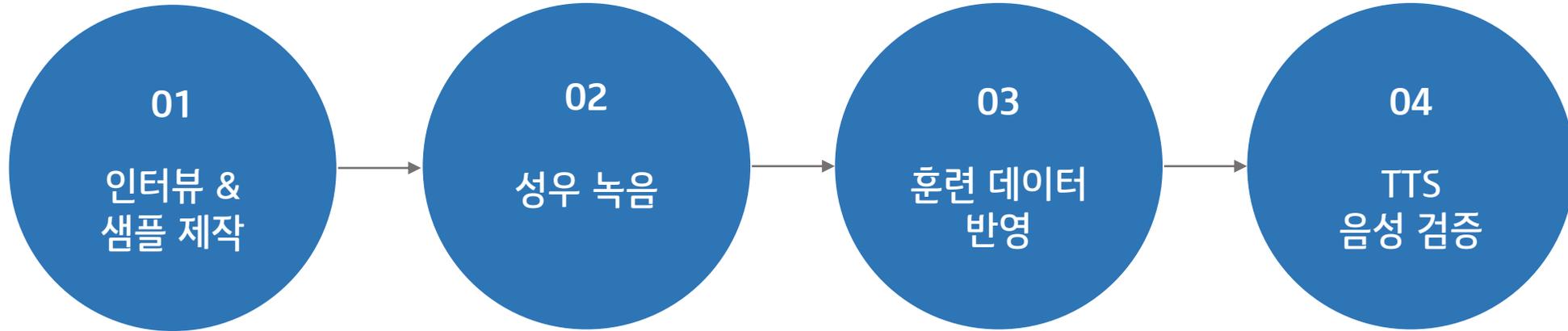


- 녹음 문장 수: 10,000 ~ 20,000 문장
- 녹음 소요 시간: 100 ~ 200시간
- 개발 비용: 약 1억원
- 개발 기간: 약 6~8개월

※ 기본 TTS 보이스의 합성음(세기, 속도)은 바로 적용 가능
 - 속도(빠르게, 느리게), 세기(부드럽게, 강하게)

※ 다솜 목소리의 감정(3종), 감성(1종)의 경우 개발 완료된 합성음으로, 추가 개발 비용 없음
 ※ 고객사 당 1종의 보이스 스타일 개발 무료 지원

별첨 고객사 보이스 스타일 개발 프로세스



- 고객 인터뷰 내용을 바탕으로 TTS 샘플 제작
- 필요 시 사용자 인터뷰 시행
- 의견 조율 후 스타일 확정

- 서비스 관련 문장/단어 수집
- 성우 녹음 가이드 제작
- 녹음 스크립트 제작
- 성우 녹음

- 발음열 변환 작업
- 데이터 정제
- 딥러닝 훈련

- 발화 검증
- 추가 녹음 및 훈련
- 시스템 적용

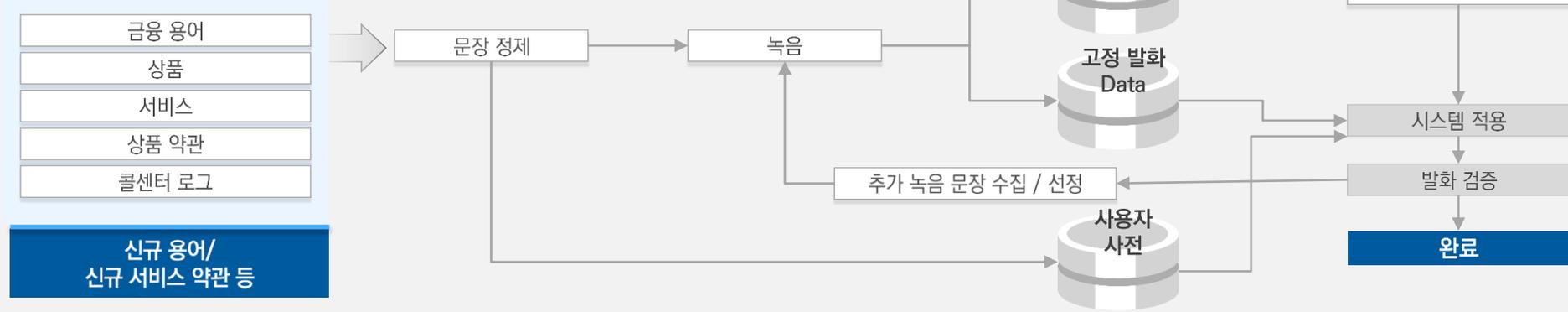
별첨 운영관리

새로운 영역의 서비스가 추가될 경우, 관련 문장을 수집하여 추가 녹음 및 학습하여 TTS 시스템 반영하도록 되어 있습니다.

● TTS 운영관리 프로세스

※ 금융용어 TTS 모델 학습 및 고도화

- 적용절차: 고객 요청 > **KT 연구소 개발**(추가 녹음, 학습, 검증 등) > 고객사 적용
- 새로운 영역의 서비스가 추가될 경우, 관련 문장을 수집하여 추가 녹음 및 학습하여 TTS 시스템에 반영
- 발화 필요한 문장에 특수 문자, 기호, 일반 발화와 다른 발음 문구의 경우 사용자 사전 등록

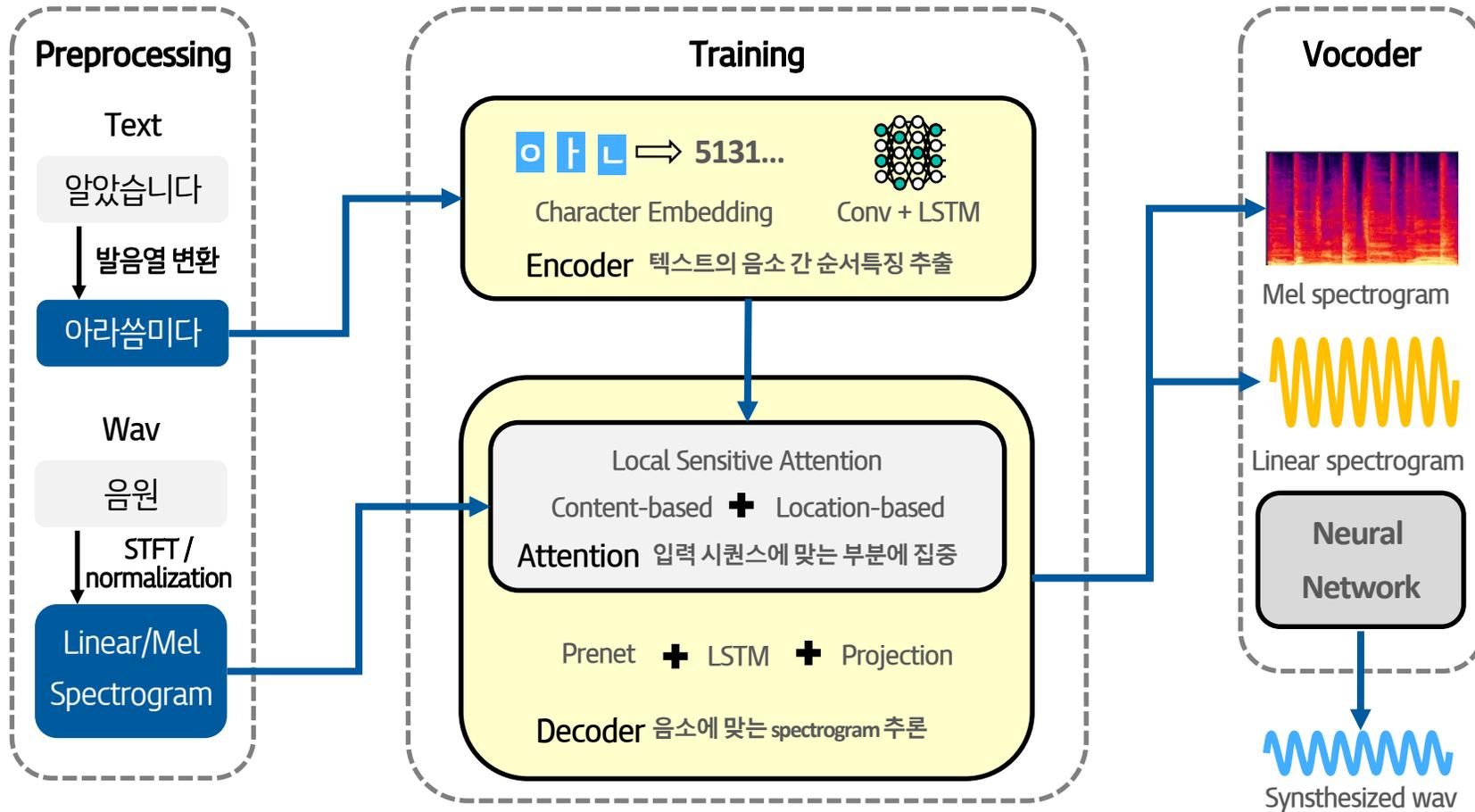


※ TTS 추가 개발 프로세스



별첨 용기원 딥러닝 기반 P-TTS 엔진

구글 Tacotron을 기반으로, GAN Vocoder를 결합한 자체 개발 엔진을 보유하고 있습니다.



별첨 TTS 솔루션 하드웨어(HW) 권장 스펙

300석 기준으로 정리된 HW 권장스펙, 서비스 시나리오/발화 내용 따라 변경될 수 있음

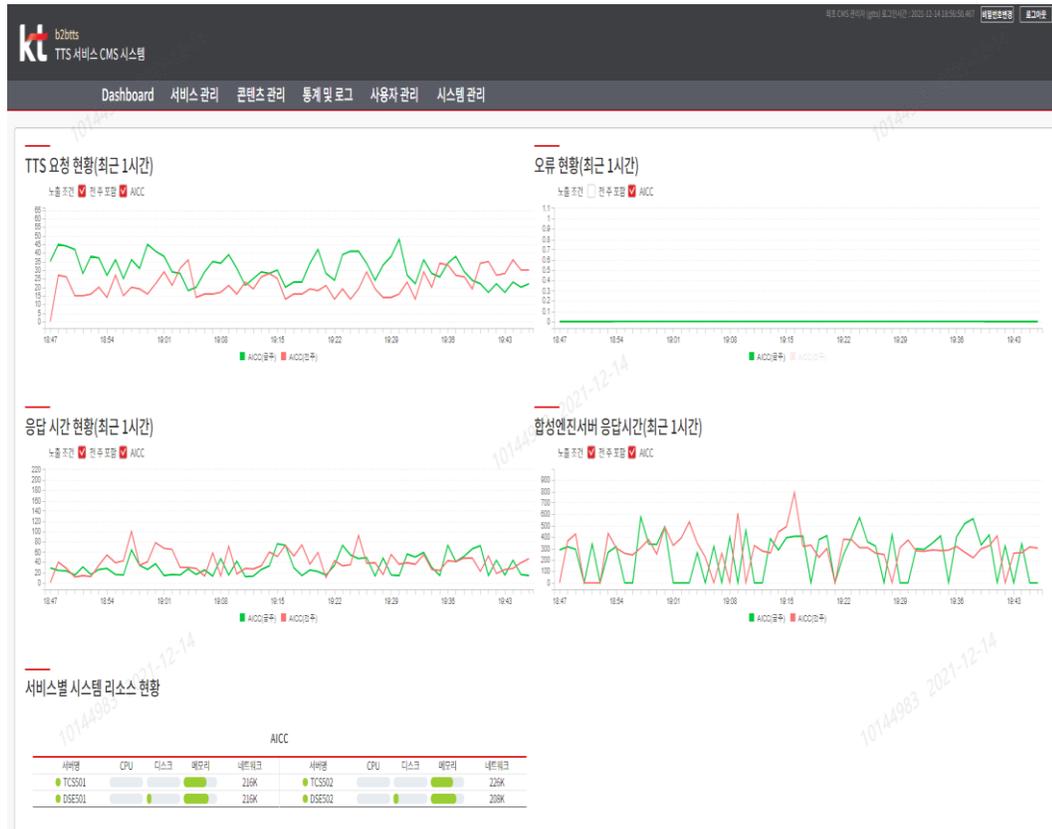
서버명	상세스펙(내용)	수량	비고
TCS	CPU: 16 core, 2.7GHz, Mem:64G, Disk: 200GB	1 (2)	이중화 적용시 2대
DSE (CPU)	CPU: 32core, 2.7GHz, Mem:64G, Disk: 200GB	5	
CMS	CPU: 8 core, 2.7GHz, Mem:16G, Disk: 100GB	1	TTS 관리/모니터링
DB	CPU: 8 core, 2.7GHz, Mem:16G, Disk: 100GB	1	TTS 이력 및 관리 데이터 저장
합계		5 (6)	

별첨 KT 자체 특허출원 및 기술 보유(11건, BM특허 및 기술특허)

KT 자체 특허출원 및 기술 보유(11건, BM특허 및 기술특허)		
구분	설명	출원일
BM 특허	가족 정보에 기초하여 음성 데이터를 생성하는 장치, 방법 및 컴퓨터 프로그램	2020/11/23
	동화 낭독 서비스를 제공하는 서버, 방법 및 컴퓨터 프로그램	2019/08/12
	감정 음성을 합성하는 장치, 방법 및 컴퓨터 프로그램	2019/09/16
	이미지를 통한 학습을 이용하여 합성 음원을 생성하는 장치, 방법 및 컴퓨터 프로그램	2019/08/27
기술 특허	음성 합성 스케줄을 조정하는 캐쉬 서버, 방법 및 음성 합성을 수행하는 음성 합성 서버	2020/10/13
	음성 합성 서비스를 제공하는 서버, 단말 및 방법	2019/07/30
	TTS 서비스 제공 방법 및 TTS 서비스 제공 시스템	2019/11/04
	다중 채널을 통해 멀티미디어 데이터를 전송하는 사용자 단말, 서버 및 방법	2019/09/27
	음성 합성 방법 및 음성 합성 장치	2018/10/18
	TTS 서비스 제공 방법 및 이를 위한 중계 서버	2018/11/01
	음성 합성 서비스를 제공하는 서버, 사용자 단말 및 방법	2018/08/02

별첨 CMS UI

- ✓ 대시보드 : 요청 통계, 응답시간, 오류현황 등에 대한 최근 1시간동안의 차트 제공
- ✓ 상용구 : 상용구 개별 / 대량 등록 및 배포
- ✓ 사용자사전 : 사용자사전 개별 / 대량 등록 및 배포



사용자 사전 목록 조회

사건유형: 전체 | 상태: 전체

등록일: 2018-01-01 ~ 2021-12-14

페이지 개수: 10 | 전체검색결과: 3건

번호	사건유형	시	검색어	지판번호	등록일	최종 배포일	상태
1	사용자사전	시	에이아이		2021-08-30 15:37:58	2021-08-30 15:37:58	배포완료
2	사용자사전	문의	무니		2021-08-13 14:10:48	2021-08-13 14:10:48	배포완료
3	사용자사전	코로나19	코로나19		2021-08-13 09:59:06	2021-08-13 09:59:06	배포완료

Excel 다운로드 | 초기화 | 등록

상용구 목록 조회

등록일: 2018-01-01 ~ 2021-12-14

서비스스피커: 전체 | 상태: 전체

파일명: | 검색어: | 상용구: | 검색키워드: |

페이지 개수: 10 | 전체검색결과: 225건

번호	서비스	상용구	스피커	샘플레이트	등록일	배포일	상태
1	AICC	음성 제1 단계드라마?	LDS_GP	24 kHz	2021-11-03 16:01:38	2021-11-03 16:01:38	배포완료
2	AICC	저희 예약은 연중 무휴로 운영됩니다.	LDS_GP	24 kHz	2021-11-03 11:35:13	2021-11-03 15:57:12	배포완료
3	AICC	네 고객센터, 예약이 많으셔서 취소는 확실히 통보해 드립니다.	LDS_GP	24 kHz	2021-11-03 11:35:13	2021-11-03 11:35:14	배포완료
4	AICC	예약이 많으셔서 취소는 확실히 통보해 드립니다.	LDS_GP	24 kHz	2021-11-03 11:35:13	2021-11-03 11:35:14	배포완료
5	AICC	발행 시간을 말씀해주세요.	LDS_GP	24 kHz	2021-11-03 11:35:13	2021-11-03 11:35:13	배포완료
6	AICC	최종에 드릴까요?	LDS_GP	24 kHz	2021-11-03 11:35:13	2021-11-03 11:35:13	배포완료
7	AICC	첫번째와 두번째 예약 중 어떤것을 변경해 드릴까요?	LDS_GP	24 kHz	2021-11-03 11:35:13	2021-11-03 11:35:13	배포완료
8	AICC	변경해 드릴까요?	LDS_GP	24 kHz	2021-11-03 11:35:13	2021-11-03 11:35:13	배포완료
9	AICC	예약이 많으셔서 변경은 확실히 통보해 드립니다.	LDS_GP	24 kHz	2021-11-03 11:35:13	2021-11-03 11:35:13	배포완료
10	AICC	예약 변경이 취소되었습니다.	LDS_GP	24 kHz	2021-11-03 11:35:13	2021-11-03 11:35:13	배포완료

Excel 다운로드 | 초기화

DIGICO **KT**