



KT AICC 솔루션 STT(Speech to Text) 소개자료

목차

01 도입 배경

02 솔루션 개요

03 적용 기대효과

04 솔루션 적용 예시

05 솔루션 특징점

06 솔루션 주요기능

07 솔루션 요금

#. 별첨

01. 도입 배경

상담업무가 상담사의 경험과 업무 지식에 의존도가 높으며, 상담 녹취 데이터의 활용도가 낮아 상담업무의 품질 관리 향상에 한계가 있습니다.

✓ 상담사의 경험과 업무 지식에 의존, 솔루션 품질 차이



상담사가 고객에게 빠르게 응대를 할 수 있도록
상담사에 대한 교육에 많은 비용을 투자,
상담사의 잦은 이직 및 총원의 어려움

✓ 상담 녹취 데이터 축적, 고객의 소리 파악에 활용 저조

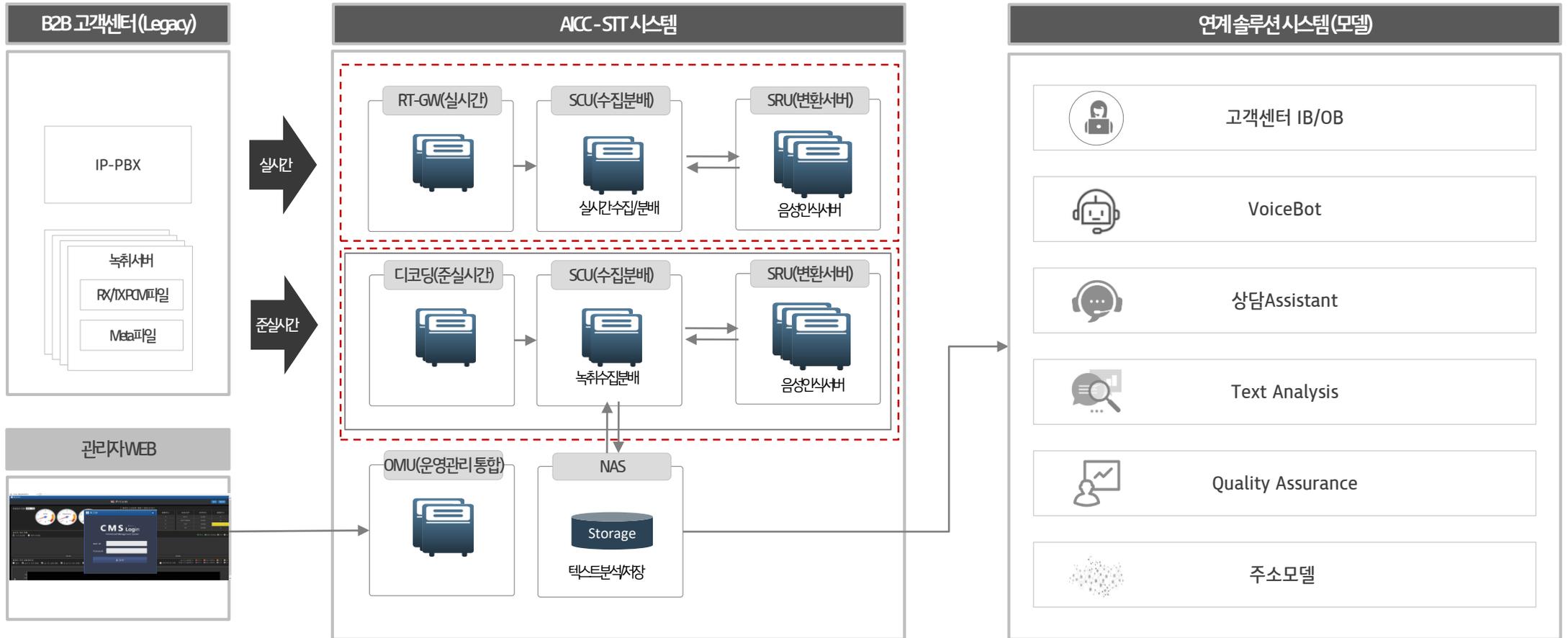


민원처리에 관한 법률에 근거하여 상담내용을 녹취,
녹취 데이터 활용 저조로
민원 소지가 많은 상품에 대한 사전 인지의 어려움

고객과 실시간으로 양방향 커뮤니케이션이 가능한 STT(Speech to Text) 기술 요구
고객센터 녹취 자료의 활용을 통한 부가가치 창출

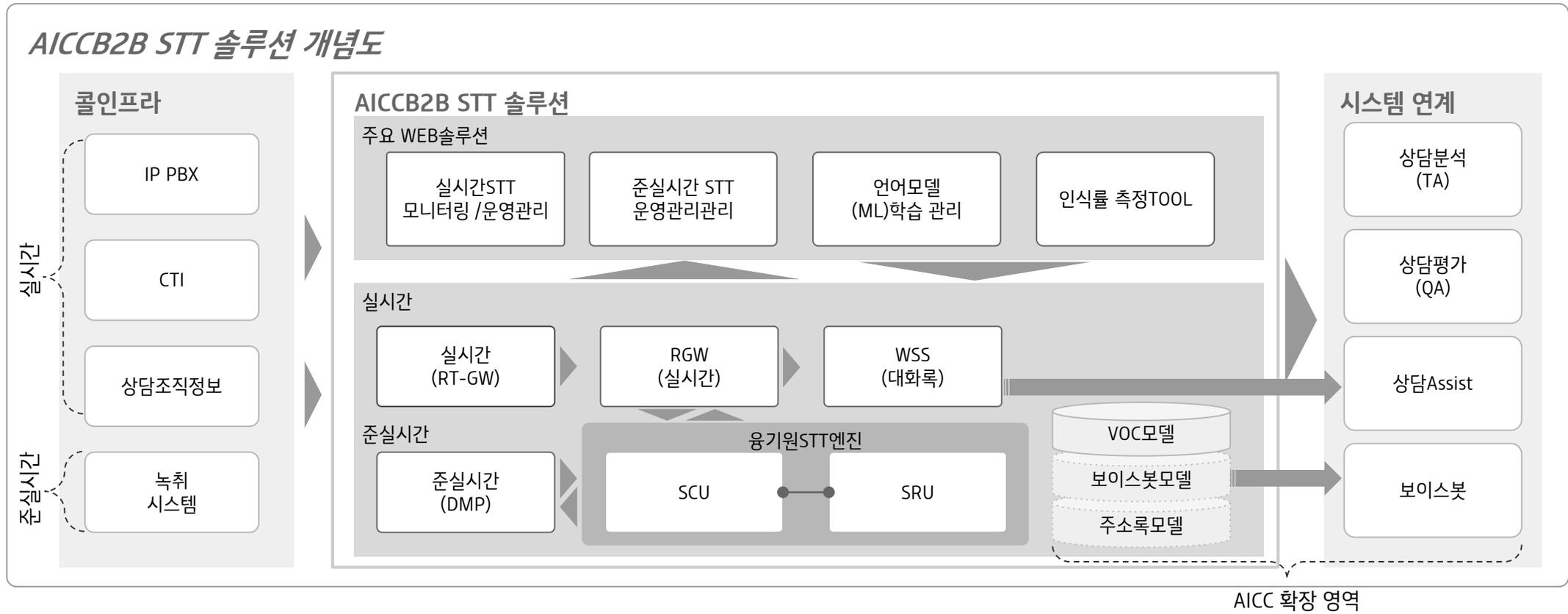
02 솔루션 개요 (개요)

AICC STT 솔루션은 실시간, 준실시간, 배치 3가지 시점에 고객의 음성을 텍스트로 변환하며, 솔루션 모델로는 고객센터IB/OB 모델, 주소모델, 보이스봇모델, 업무Assist모델이 있습니다.



02 솔루션 개요 (솔루션 흐름)

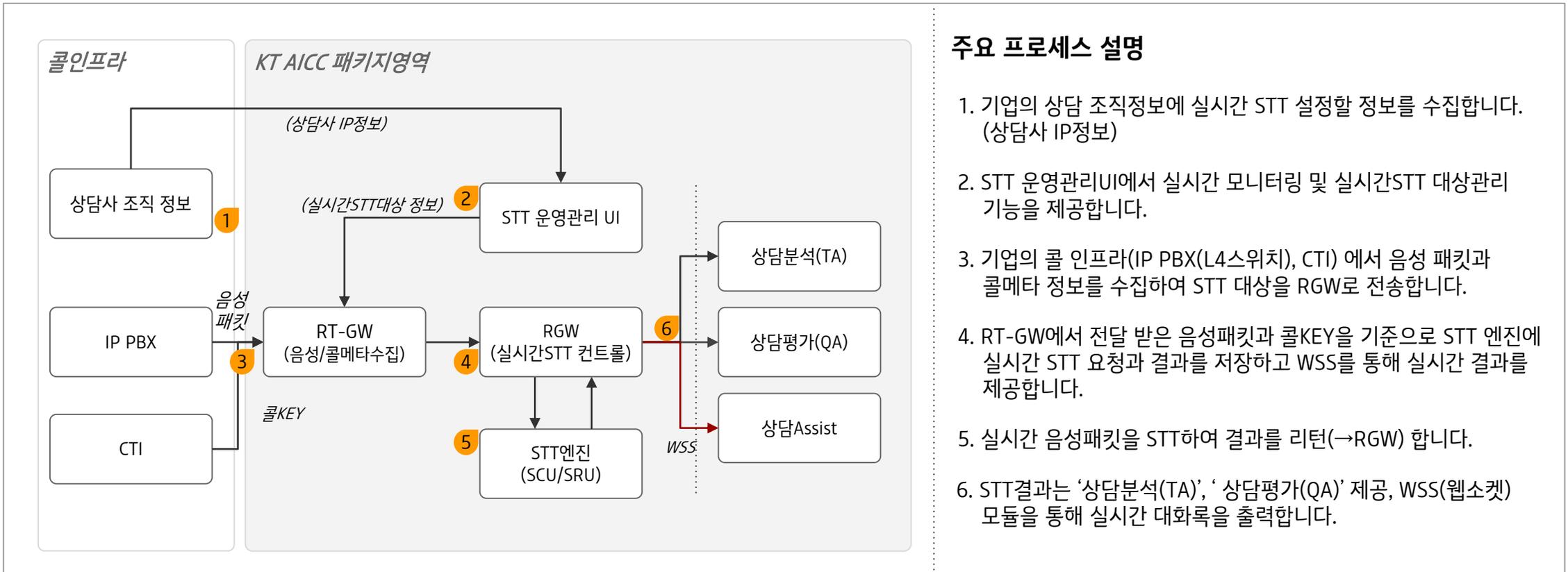
AICC STT 솔루션은 기업의 콜인프라 환경과 연계하여 실시간 / 준실시간, 언어모델 학습, 인식률 측정 등 환경을 제공하고 STT된 결과는 상담분석(TA), 상담평가(QA), 상담Assist 등 에 결과를 활용할 수 있도록 구성되어 있습니다.



02 솔루션 개요 (실시간STT - 처리 프로세스)

고객센터에서 실시간 STT 대상을 선택적 적용/운영 할 수 있도록 설정, 고객센터 자원을 단계적으로 적용할 수 있는 환경을 제공합니다.

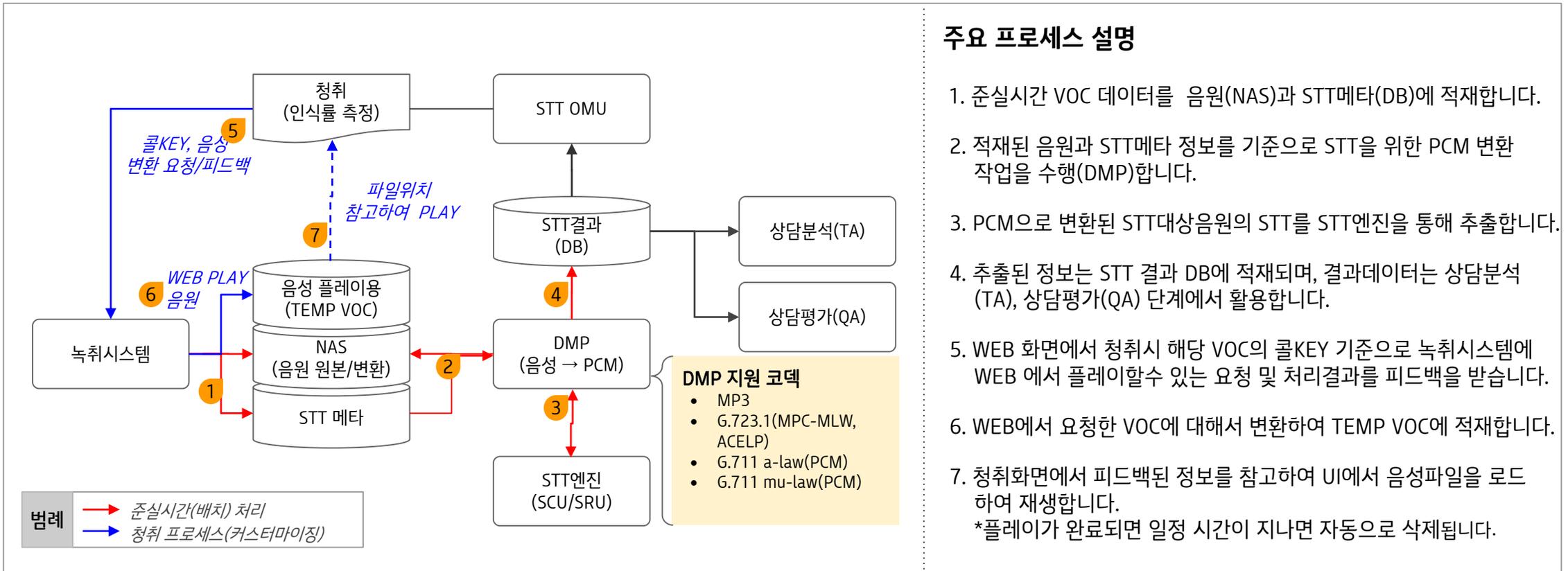
실시간 STT 처리 프로세스



02 솔루션 개요 (준실시간STT - 처리 프로세스)

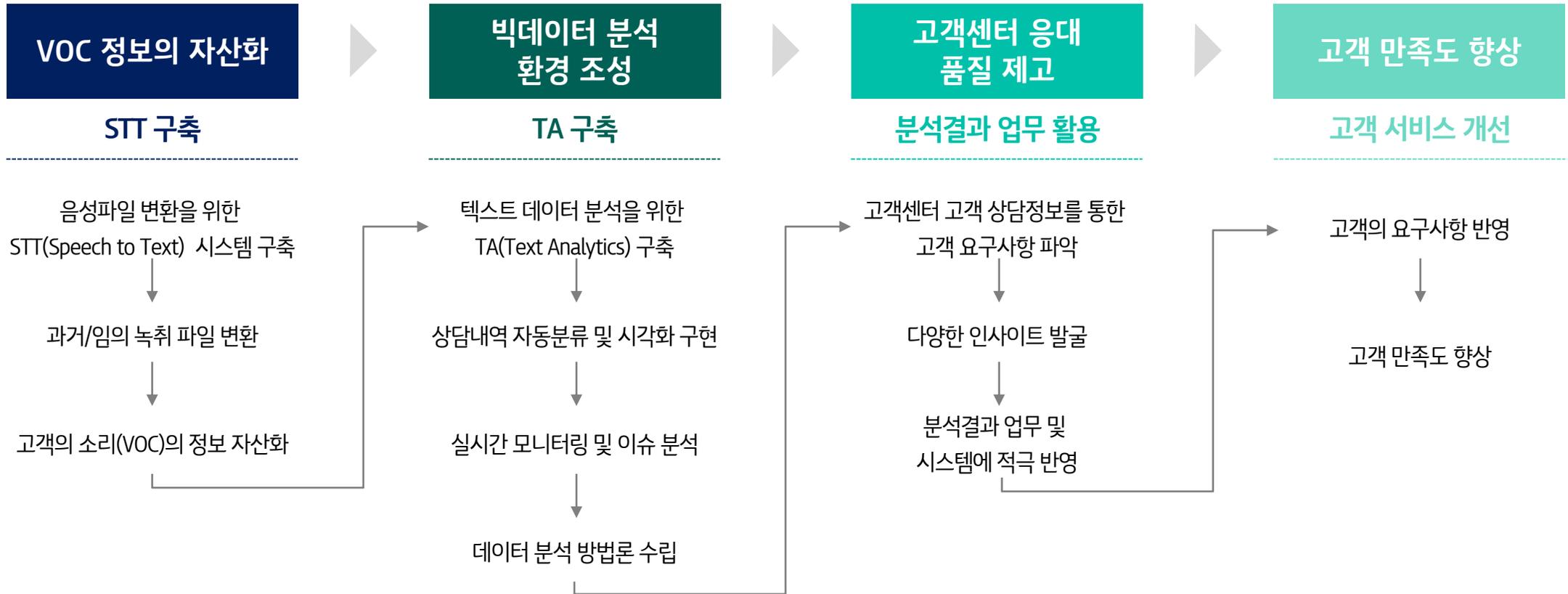
기업의 녹취 시스템과 연계하여 고객센터의 VoC데이터를 텍스트로 변환하여 DB에 적재하여 상담분석(TA), 상담평가(QA)로 결과를 제공합니다.

준실시간 STT 처리 프로세스



03 적용 기대 효과

STT(Speech to Text) 시스템 구축을 통하여 VoC(고객의소리) 데이터의 정보 자산화와 빅데이터 분석을 위한 환경 조성을 기반으로 업무 및 시스템에 분석결과를 활용함으로써 고객센터 응대의 품질 제고와 고객만족도 향상에 기여합니다.



04 솔루션 적용 예시

‘18년 4Q에 착수하여 기한내 완수, AI기반 자동분류 최초적용, 실시간 상담 어드바이저 성공적 수행 완료했습니다.

라이나생명 STT/TA 구축

• 고객상담 녹취를 STT변환하여 자동으로 분류/분석하여 상담품질을 관리하고 서비스 개선을 도모하는 시스템



- 사업명 : 라이나생명 용산 CS센터 STT/TA 시스템 구축
- 구축기간 : '19.10 ~ '20.04 (6개월)

구분	과업	목표기준	
STT	• STT 솔루션 도입 및 음성데이터 자산화를 위한 DB 구축	목표 인식률 85%	청취처리율 감소

하나캐피탈 STT/TA 구축

• 실시간 STT분석하여 QA자동평가 / TA 분석을 자동으로 진행하는 시스템



- 사업명 : 하나캐피탈 STT/TA 시스템 구축
- 구축기간 : '19.9.18 ~ '20.03.17 (6개월) (20.02.17 오픈, 안정화 1개월)

구분	과업	목표기준	
STT	• 실시간 STT 솔루션 도입 및 음성데이터 자산화를 위한 DB 구축	상담사 92% / 고객 82%	청취처리율 감소

05 솔루션 특징점

검증된 KT 융기원 기술기반의 디바이스와 솔루션 엔진으로 STT(Speech to Text) 최적화 솔루션을 구현하고 GUI 기반 관리 툴 강화로 패키지의 완성도를 높였습니다.

고객 니즈에 맞춘 음성모델 제공

- 콜 + 컨트롤 데이터: 보이스 & 데이터(컨트롤) 결합
- 복수 대표번호 / 번호별 다양한 시나리오 운영

발화기준 음성 barge-in 기능 제공

- IP-PBX에 단말(소프트폰)으로 연동
- SIP 프로토콜을 이용한 콜 컨트롤

고객 상담 도메인 음성 트레이닝

- 대화처리와 연계한 보이스봇 전용 음성모델 학습
- 언어모델(LM) 학습 관리환경 제공

표준 기술 및 최신 트렌드 기술 적용

- JAVA/JSP/Non Active X 기반의 프로그램 및 웹 표준 기반 멀티브라우저 지원(IE10 이상 Non Active X 지원)
- IT 운용환경에 원활한 연계를 위한 전자정부 표준 프레임워크 기반 설계

우수한 한국어 음성인식률 제공

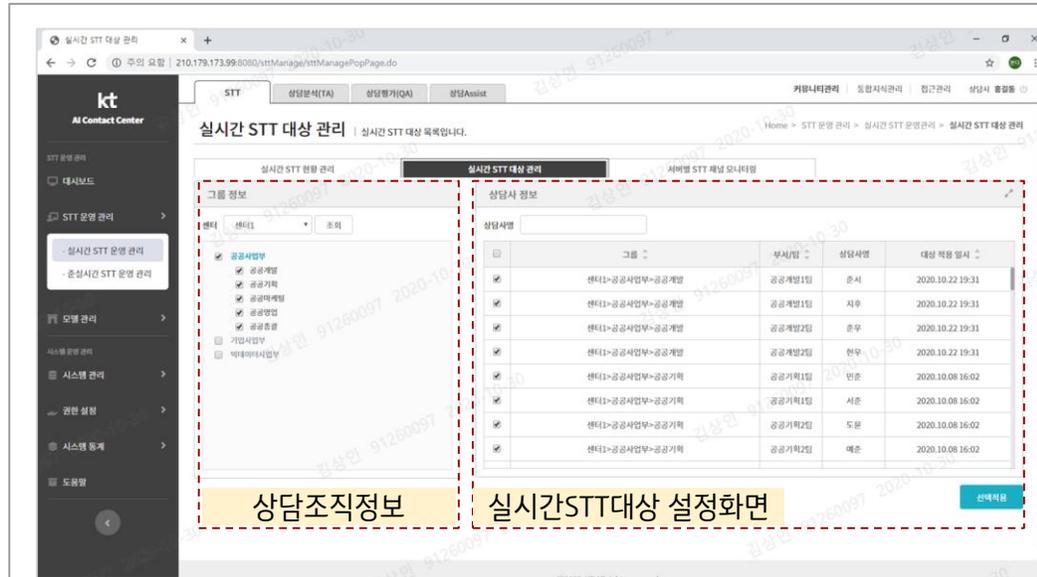
- 국내환경에 맞는 음성인식 원천 기술 보유로 고객 맞춤형 엔진 제공
- 빠르고 정확한 실시간 스트리밍 등 음성인식 기술 확보

06 솔루션 주요 기능 (실시간STT - 운영관리 기능(UI))

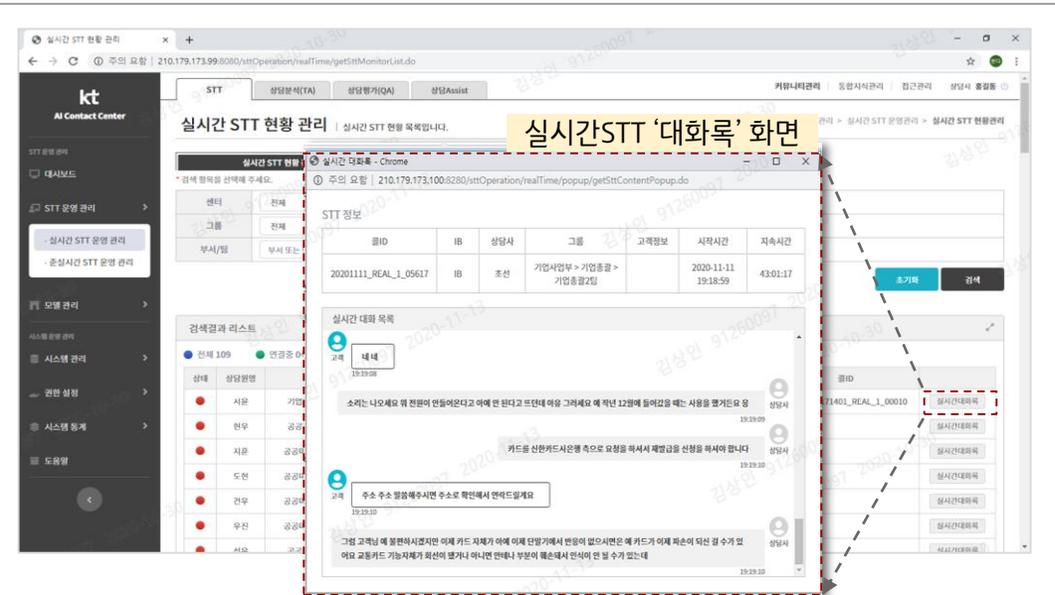
실시간 STT로 처리할 대상을 선택적으로 적용 운영할 수 있는 환경을 지원함으로써 고객센터의 자원을 효율적으로 운영할 수 있도록 합니다.

기업의 상담조직과 연계하여 ‘실시간STT’ 적용 기능 제공

실시간STT 모니터링



기업의 상담조직과 연계하여 ‘실시간 STT 대상’ 데이터를 수집(배치)하여 STT솔루션 관리 화면에서 ‘실시간STT’ 대상 관리기능을 통해 관리 운영할 수 있도록 제공합니다.

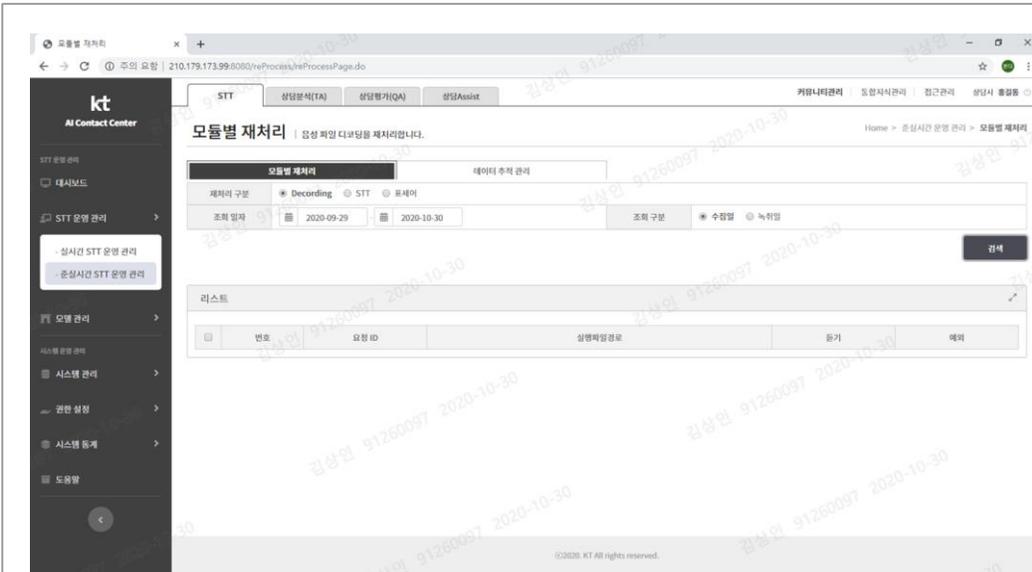


실시간 STT 모니터 화면으로 상담사IP 기준으로 통화중 상태를 출력하고 ‘실시간 대화록’ 기능을 통해 실시간 STT 결과를 문장 단위로 출력하여 실시간 음성을 텍스트로 변환하는 과정을 확인할 수 있습니다.

06 솔루션 주요 기능 (준실시간 STT - 운영관리 환경 제공)

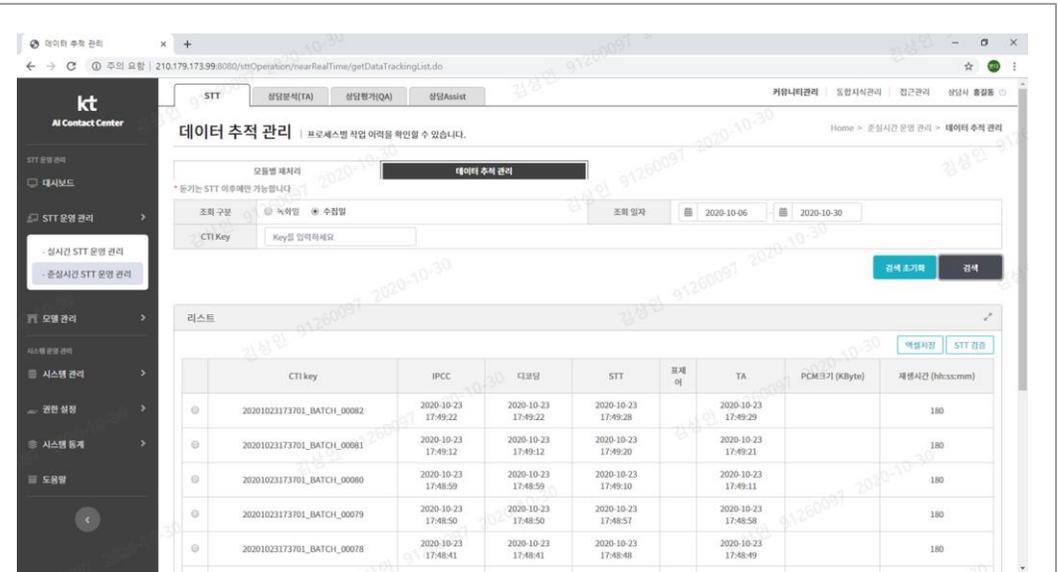
준실시간 STT 결과 모듈별 재처리, 데이터 추적 관리 등 후처리 강화를 통해 준실시간 STT 의 효율적 관리/운영 환경을 제공합니다.

모듈별 재처리



준실시간 STT 운영 관리 기능으로 준실시간 처리(배치) 중 오류가 발생 시 솔루션에서 제공하는 모듈 재처리 기능을 통해 손쉽게 운영할 수 있는 지원 기능이 구성되어 있습니다.

데이터 추적 관리



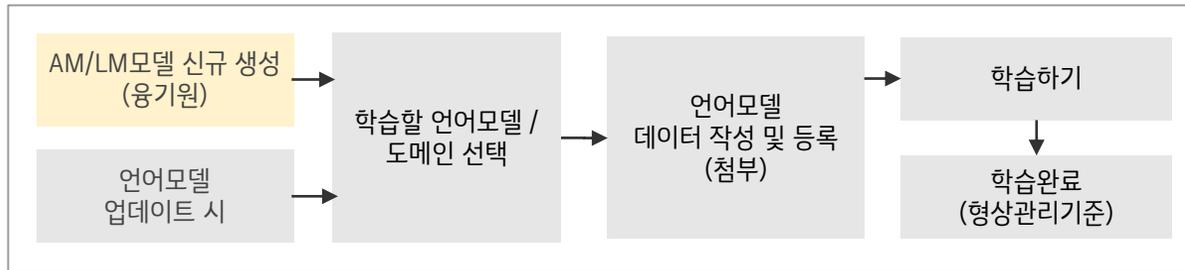
준실시간 STT 운영 관리 기능으로 수집단계부터 단계적으로 처리 현황을 모니터링할 수 있으며, 해당 VOC STT를 결과와 청취할 수 있는 기능이 제공됩니다.

06 솔루션 주요 기능 (WEB 환경의 언어모델(LM) 관리 기능 제공)

기존 용기원에서 CMD 환경에서 언어모델(LM) 학습 스크립트를 AICCB2B사업을 통해 WEB환경에서 사용자가 언어모델 및 도메인 별 언어모델(LM)을 학습하고 관리할 수 있는 UI 환경 제공합니다.

TO-BE (WEB 환경 언어모델 학습/관리 체계)

1. WEB환경의 언어모델 관리 환경 체계 구성



<언어모델 학습 프로세스>

2. WEB환경에서의 언어모델별, 도메인 (VOC언어모델, 보이스봇 언어모델, 주소록 언어모델 등)별 멀티모델 관리 기능 제공
3. WEB 환경에서 언어모델(LM)학습 데이터 작성하여 학습할 수 있는 환경 제공



<WEB 환경의 언어모델별, 도메인별 관리 화면>



<WEB 환경의 언어모델(LM) 학습 화면>

06 솔루션 주요 기능 (WEB 환경의 인식률 측정 TOOL 적용)

기존 인식률 측정 방법은 단계가 복잡(숙련자용)하였으나, 사용자가 손쉽게 인식률을 측정하고 인식률 추이 분석 및 품질 관리를 할 수 있도록 패키지가 구성되어 있습니다.

TO-BE (WEB 환경 인식률 측정 및 관리 체계)

1. WEB환경의 인식률 측정 TOOL 환경 제공



<인식률 측정 프로세스>

2. 사용자가 WEB환경에서 STT된 결과와 음성플레이 환경을 통한 인식률 측정을 위한 전사 작업 환경 지원
3. 인식률 TOOL로 측정된 결과 추이 리포트를 통해 인식률 품질 관리 환경을 지원

AICCB2B STT패키지 개발 화면

<인식률 추이 차트>

음성플레이

STT 결과	전사자료 (텍스트 입력)	인식률	인식률 결과
내	남이 짝는 놔 이 소름입니다	100.00 %	STT 글자 수 136 글자
내	남이 짝는 놔 이 소름입니다		정답 글자 수 136 글자
내	아 그제세요 인종해우신 내		오답 글자 수 0 글자
내	아 그제세요 인종해우신 내		삽입 0 글자
내	아 네 인종해우신 이 강선 고재님 본인입니까		삭제 0 글자
내	아 네 인종해우신 이 강선 고재님 본인입니까		대사
내	네 알에 새년월일까지 알림해주시겠습니까		
내	네 알에 새년월일까지 알림해주시겠습니까		
내	네 알에 새년월일까지 알림해주시겠습니까		

<WEB 환경의 인식률 측정TOOL>

07 솔루션 요금 (실시간STT)

도입 비용은 솔루션에 대한 라이선스 비용 + 개발비(커스터마이징) + 인프라 비용으로 구성되며 라이선스 외 비용은 고객사 환경, 요구수준에 따라 상이하므로 별도 협의가 필요합니다.

구분	요금 (VAT 별도)	상세 내역
라이선스 비용	200,000,000원/1식 (STT 서버) 500,000/1CH (STT 채널)	<ul style="list-style-type: none"> ▪ 솔루션 라이선스에 대한 List Price <ul style="list-style-type: none"> - STT 서버 라이선스 : K3-Voice Server (RT Sever License 포함) - STT 채널 라이선스 : K3-Voice Channel
개발비	(별도 견적)	<ul style="list-style-type: none"> ▪ STT 실시간 서버 환경 구축 (RT-GW 설치 및 구동) ▪ 콜인프라 분석, 설계 및 연동 ▪ STT 모델 구축 (음향/언어모델 학습 및 튜닝) ▪ 전사작업 (학습 및 검증 비용 포함)
인프라 비용 (H/W)	(별도 견적)	<ul style="list-style-type: none"> ▪ 최소구성 : RT-GW(1) / SCU(1) / SRU(2) / OMU(1) / DB(1) / Storage(1) ▪ 데이터 규모 및 요구 성능에 따라 상이 ▪ 고객사 환경 및 요건 협의 후 정확한 구성/비용 산정 가능

07 솔루션 요금 (준실시간STT)

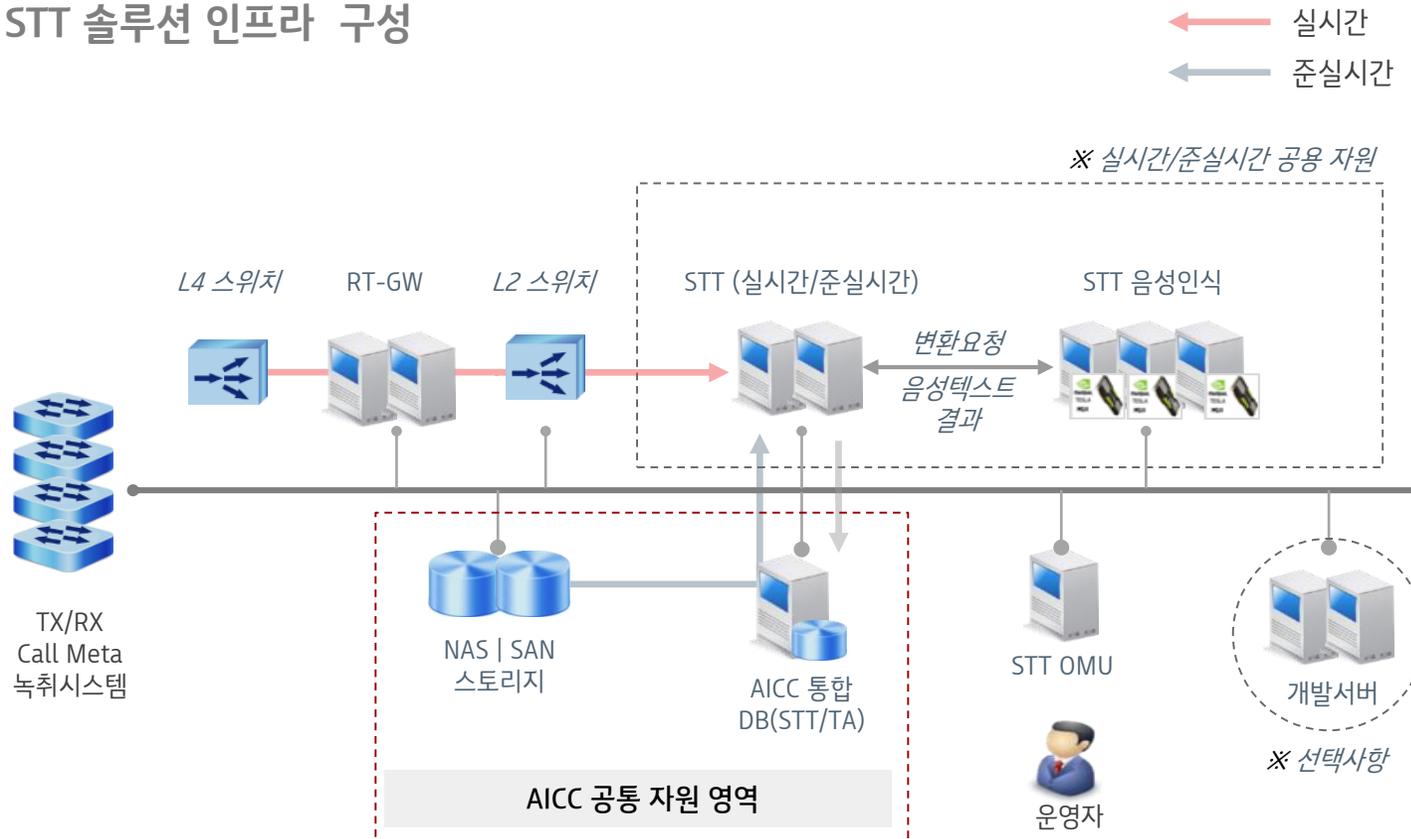
도입 비용은 솔루션에 대한 라이선스 비용 + 개발비(커스터마이징) + 인프라 비용으로 구성되며 라이선스 외 비용은 고객사 환경, 요구수준에 따라 상이하므로 별도 협의가 필요합니다.

구분	요금 (VAT 별도)	상세 내역
라이선스 비용	200,000,000원/1식 (STT 기본 서버) 10,000,000/1식 (STT 부가 서버)	<ul style="list-style-type: none"> ▪ 솔루션 라이선스에 대한 List Price <ul style="list-style-type: none"> - STT 기본 서버 라이선스 : K3-Voice Server license - STT 부가 서버 라이선스 : STT SRU 서버 라이선스
개발비	(별도 견적)	<ul style="list-style-type: none"> ▪ STT 준실시간 서버 환경 구축 (STT 엔진 설치 및 구동) ▪ 콜인프라 분석, 설계 및 연동 ▪ STT 모델 구축 (음향/언어모델 학습 및 튜닝) ▪ 전사작업 (학습 및 검증 비용 포함)
인프라 비용 (H/W)	(별도 견적)	<ul style="list-style-type: none"> ▪ 최소구성 : SCU(1) / SRU(2) / OMU(1) / DB(1) / Storage(1) ▪ 데이터 규모 및 요구 성능에 따라 상이 ▪ 고객사 환경 및 요건 협의 후 정확한 구성/비용 산정 가능

A 시스템 구성 (솔루션 인프라(N/W) 환경(최소 사양))

AICC STT 솔루션은 기업의 인프라 환경을 고려하여 설계되어 있으며, 실시간과 준실시간 STT을 위한 환경을 다음과 같이 구성하였습니다.

STT 솔루션 인프라 구성



Description

- 음원수집은 녹취서버 연동을 위해 NAS 서버를 구성
- RT-GW 는 L2 스위치와 물리적으로 NIC 연결 가능 공간에 위치
- SCU 이중화를 위해 L4 스위치 구성
- SRU 서버는 자체 이중화 구성 (N + 1)

<참고사항>
 기본적으로 구성으로 기업의 채널수에 따라 인프라 환경 구성 필요

A 시스템 구성 (소프트웨어(SW) 구성도)

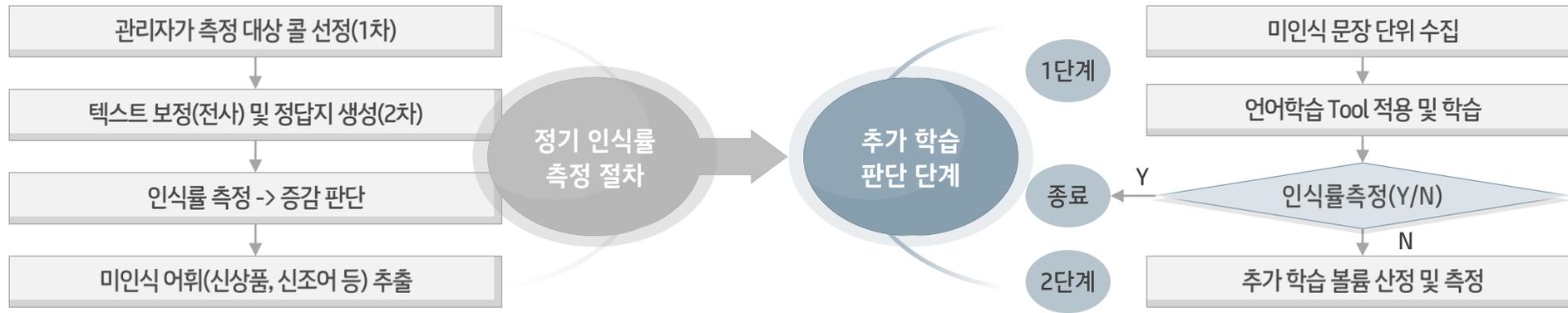
AICC STT 솔루션의 소프트웨어 구성으로 기업의 300석 기준으로 정리된 자료이며, 기업의 인프라 환경에 따라 변경될 수 있습니다.

	RT-GW	SCU	SRU	OMU
준실시간 STT SW		<ul style="list-style-type: none"> SFP (녹취 분리/분배) DMP (녹취 조회/변환) 	<ul style="list-style-type: none"> KFP - 인식서버 (Front End Process) 	<ul style="list-style-type: none"> STT관리운영
실시간 STT SW	<ul style="list-style-type: none"> SCS (콜정보수집) VAS (PCM변환/전송) PFS(패킷필터링) 	<ul style="list-style-type: none"> DIP (세션 관리 프로세스) GSP (음성 스트림 연동) RGW(STT-GW 연동) WSS(상담 대화록 연동) 	<ul style="list-style-type: none"> KEP - 인식서버 (Back End Process) STT 엔진 모델 Kaldi Lib GPU/cudaLib 	<ul style="list-style-type: none"> OMP KT-WCR LM-Tools Apache 2.4 Tomcat 8.5
공통	<ul style="list-style-type: none"> CentOS 7.6 	<ul style="list-style-type: none"> SRP(서비스 관리) Kafka 1.0,1 GENIFP(운영시스템연동) SYSM(시스템 리소스 연동) PNR(프로세스 관리) CentOS 7.6 	<ul style="list-style-type: none"> GENIFP(운영시스템연동) SYSM(시스템 리소스 연동) PNR(프로세스 관리) CentOS 7.6 	<ul style="list-style-type: none"> GENIFP(운영시스템연동) SYSM(시스템 리소스 연동) PNR(프로세스 관리) CentOS 7.6

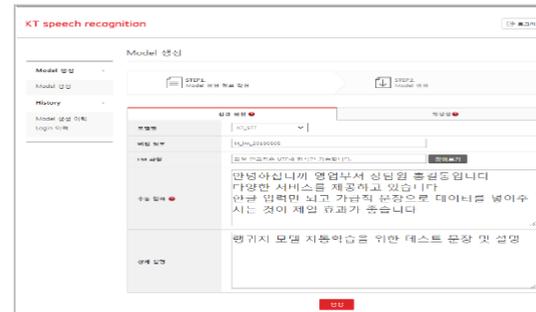
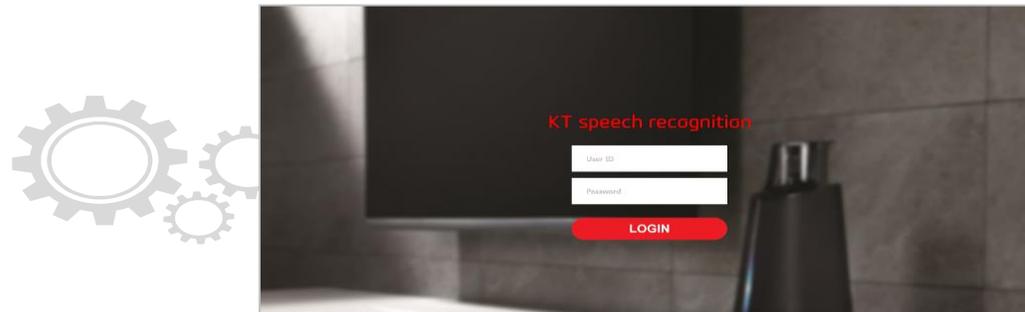
A STT(Speech To Text)-인식률 측정

음성 인식률 측정 절차를 활용하여 인식률 유지 여부를 판단하고 추가학습 필요를 인지할 수 있는 단계를 적용, 언어학습 자동 Tool을 제공하여 신속하게 인식률을 향상시킵니다.

정기 인식률 측정 절차



언어모델 학습 Tool



B FAQ

Q) 개인정보는 STT에서 어떻게 처리되나요?

A) 마스킹 처리 기능이 모듈화 되어 있어 개인정보, 금융정보 등 민감정보에 대해 마스킹 처리됩니다.

Q) STT 인식률 관리는 어떻게 하나요?

A) 인식률 측정 및 모니터링 기능이 있어 STT 인식률을 자동으로 측정하고 모델별로 진단이 가능합니다.

Q) STT 인식률 제고를 위한 학습은 어떻게 관리하나요?

A) STT 학습관리는 음향모델과 언어모델 별로 진행하실 수 있고 버전관리 기능이 제공되고 있습니다.

Q) 관리 Tool 은 어떤 기능들이 제공되나요?

A) RT GW, STT, TA 통합 대시보드 환경 / 웹 기반 통합 관리운영 환경(UI, 대시보드, 리포팅 기능 개선) / 통합관리 환경 메뉴와 기능 개선 및 접근제어가 가능합니다.

DIGICO **KT**