

1. 기술의 기본정보

기술의 사업화유형 및 분류

구분	창업형	기술이전형
선택(✓)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
기술분류	IT <input checked="" type="checkbox"/> BT <input type="checkbox"/> NT <input type="checkbox"/> ET <input type="checkbox"/>	
10대 유망기술분야 또는 그 외 선택(✓)	디지털 인텔리전스	

기술명

국문	음성인식, 발음평가, 주제별 대화처리 기반의 외국어 말하기 학습 기술
영문	Computer-assisted language learning technology based on speech recognition, pronunciation evaluation, and topic-based dialogue processing

기술분류

구분	코드	대분류 기술명	중분류 기술명	소분류 기술명
국가과학기술 표준분류*	EE0202	정보/통신	소프트웨어	S/W 솔루션
6T분류**	010314	IT	정보처리 시스템 및 S/W	대화형 휴먼인터페이스 기술

대상기술과 연관된 지재산권 권리확보 현황

지식재산권 유형*	명 칭	국가명	출원번호 (출원일)	등록번호 (등록일)	출원인 /등록인	발명자
특허	Apparatus and method for recognizing speech	미국	15/187581 (2016.06.20)	-	ETRI	강병욱 외 4명
특허	Self-learning based dialogue apparatus and method for incremental dialogue knowledge	미국	15/405425 (2017.01.13)	-	ETRI	권오욱 외 2명
특허	영어 말하기 자동평가를 위한 운율 및 유창성 평가 기술	한국	2016-0032817 (2016.03.08.)	-	ETRI	이윤경 외 4명
특허	대화식 오류 교정 피드백을 이용한 언어학습 장치 및 그 방법	한국	10-2016-00328 (2016.03.18)	-	ETRI	이기영

□ 대상기술과 연관된 주요 논문 산출현황

학술지명	논문명	주저자	공동저자	게재년도	Vol(No)	페이지	SCI 등재여부
ETRI Journal	Language Model Adaptation Based on Topic Probability of Latent Dirichlet Allocation	전형배	이수영	2016	38		SCI
ETRI Journal	Online Blind Channel Normalization Using BPF-Based Modulation Frequency Filtering	이윤경	정호영 박전규	2016			SCI
IEICE	Combining Multiple Acoustic Models in GMM Spaces for Robust Speech Recognition	강병욱	권오욱	2016	E99-D No.3		SCle
EUROCALL 2016	Task-oriented spoken dialog system for second-language learning	권오욱	김영길 이윤근	2016			-

□ 후보기술 관련 정부과제

과제명	언어학습을 위한 자유발화형 음성대화처리 원천기술 개발		
주관부처명	미래창조과학부	연구책임자명	이윤근
중사업명	정보통신.방송 연구개발사업	세부사업명	-
주관기관명	한국전자통신연구원	연구기간	2015.3.1. ~ 2019.2.38.
연구비총액	총 15,400,000천원 (정부: 11,550,000천원, 자체 : 천원, 민간: 3,850,000천원)		
성과산출 기여도	(100)%		

2. 기술 개요 및 특징

(1) 기술성

□ 대상기술 정의 및 특성

○ 대상기술의 정의

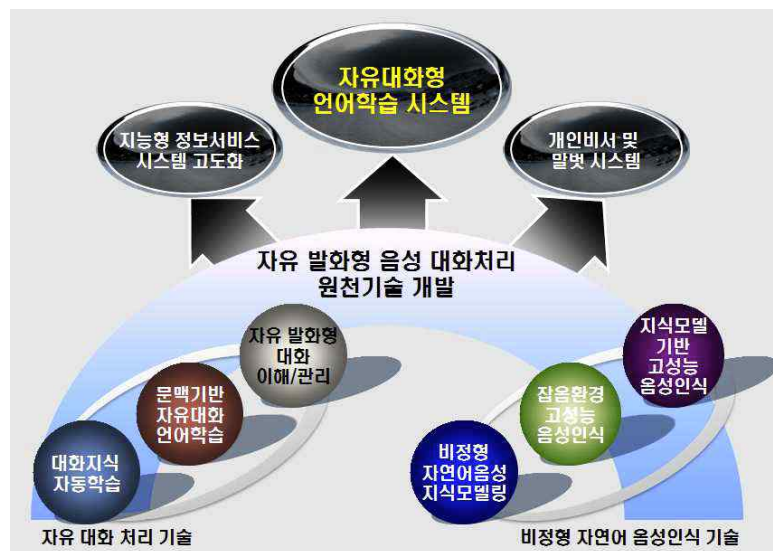
- 세계 최고 수준의 비원어민 발성 비정형 자연어 음성인식 원천 기술 개발
- 학습 주제별 대화지식 자동학습 및 자유발화형 대화처리 원천기술 개발
- 음성인식 및 자연어처리 기술에 기반한 언어학습 및 자동평가 기술 개발

○ 기존기술대비 특징점

- 자유대화형 언어학습 기술은 플랫폼 및 언어 독립적인 기술로서, 한국어/영어뿐만 아니라 중국어, 일본어, 유럽어 등 다른 언어를 위한 언어교육 시스템 및 SW에 이식 및 응용 가능함
- 음성대화 핵심원천기술은 차량 정보서비스, 콜센터 녹취 및 고객응대서비스, 노인 말벗서비스를 비롯한 개인비서서비스, 동시통역시스템 등의 성능개선 및 신규 서비스 창출, 확장이 가능함

○ 기술이 적용될 수 있는 제품 및 중요성

- 음성인터페이스 시장은 2011년부터 2017년까지 연평균 약 16.2%의 성장률을 기록, 2017년 1130억불 규모로 예상함 (HCI의 핵심 기술인 자유발화형 음성대화 처리 기술을 통해 스마트카, 지능형 로봇 등 지능형 제품 및 서비스 시장 창출 가능)
- 언어학습 세계시장 규모는 2013년 563억불이며, 특히 디지털 언어학습 시장은 아직 초기 단계로써 아시아, 미주, 유럽 등 글로벌 시장의 성장 잠재력이 아주 큼 (영어 회화 훈련 서비스 및 외국인을 위한 한국어 튜터링 서비스 등 ICT-교육 융합 신규 시장 창출 가능)



<관련 도면 및 제품 이미지>

□ 신청기술의 현재 개발 상태

해당 기술의 현재 개발 상태를 기술완성도 측면에서 작성 1) 아이디어창안 () 2) 특허만 신청·등록 () 3) 기술개발 진행 중 () 4) 기술개발완료 () 5) 시제품단계 (o) 6) 실용화단계 () 7) 시장개척단계 ()	
추가 기술개발 필요 여부	필요 (o), 불필요 ()
상용화를 위한 추가기술개발 내용	언어학습을 위한 실용화 콘텐츠, LMS 등 SW 개발
상용화를 위한 단계별 주요일정 계획	※ 추정 소요기간 및 비용포함 기재

□ 적용제품 및 응용 분야

구분	1순위	2순위	3순위
적용분야	외국어 말하기 학습	대화형 음성 인터페이스	
적용제품의 예	영어 회화 학습 S/W, 한국어 회화 학습 S/W	대화형 정보서비스 S/W	
타제품 또는 타분야 응용 가능성			

□ 기술 우수성

- o 원어민 교사 없이 컴퓨터를 이용하여 영어 말하기 연습을 수행하며, 영어 표현 학습, 문법 및 발음 교정 등을 제공받을 수 있는 음성언어기술 기반 영어 학습 서비스
- 음성인식/이해, 대화처리 기술을 말하기 학습에 적용, 컴퓨터가 원어민 교사 역할 수행
 - 학습자의 발성을 이해하고 발음, 문법 등 오류에 대한 교육적 feedback 제공
- * 기존 유사 서비스의 경우, 음성인식을 이용한 발음평가 수준이며, 피드백 기술 및 대화처리 기술 등은 기술 난이도가 높아 상용서비스에 적용되지 못하고 있음



□ 기술현황 및 전망

<p>o (국외) 음성인식 및 발음 평가 기술을 적용하여 단순한 말하기/발음 훈련 서비스가 상용화되고 있음</p> <ul style="list-style-type: none"> - 구글 벤처가 투자한 잉글리시센트럴, 로제타스톤(미국), 피어슨(영국) 등이 자체적으로 개발한 음성인식 기술을 바탕으로 자사 콘텐츠와 결합해 사업화 및 서비스 수행 중임 - TOEIC, TOEFL의 주관사인 미국의 ETS에서는 speaking test를 위한 자 동발음평가를 시험서비스 중에 있으며, 장기적으로 인간평가자(human rater)를 대체하는 컴퓨터 기반의 테스트(CBT, computer-based test) 도입을 목표로 기술을 개발 중에 있음 - 피어슨에서도 ETS와 유사하게 PTE Academic 서비스를 개발하여 말하기 평가를 시험 서비스 중임 <p>o (국내) 대화형 음성언어 기술에 기반하는 영어학습 기술은 대화 자유도가 제한된 수준의 파일럿 시스템이 개발된 상태임</p> <ul style="list-style-type: none"> - ETRI에서 개발한 대화형 영어학습 서비스인 ‘지니튜터’를 동아출판, 지앤비교육 등 영어교육전문기관에서 사업화 목표의 시범서비스를 진행 중임 - 컴퓨터가 정해진 주제에 따라 질문하고, 학습자는 문맥에 맞춰 대화를 진행하는 수준이며, 내용/문법/표현 오류에 대한 피드백 제시가 가능함

[2] 권리성

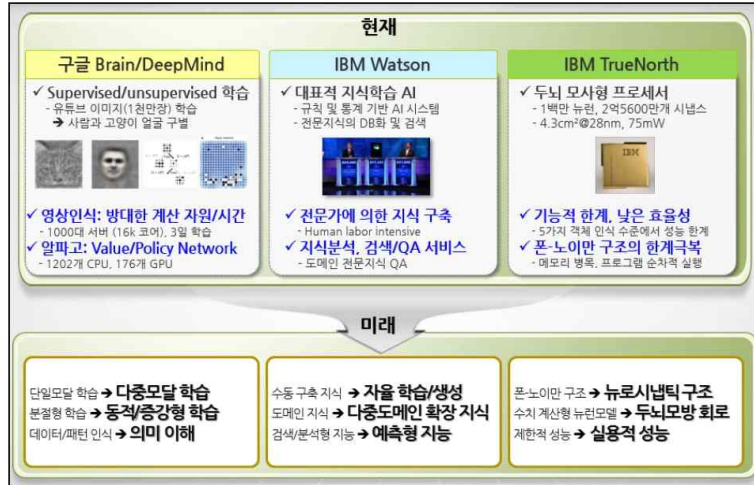
□ 핵심특허

출원(등록)번호	미국출원: 15/187581
대표청구항	A speech recognition apparatus based on a deep-neural-network(DNN) sound model, comprising: a memory; and a processor configured to execute a program stored in the memory, wherein, as the program is executed, the processor generates sound-model state sets corresponding to a plurality of pieces of set training speech data included in multi-set training speech data, generatss a multi-ser state cluster from the sound-model state sets, and sers the multi-set training speech data as an input node and the multi-set state cluster as on output node so as to learn a DNN structured parameter, and when a user' s speech and characteristic information thereof are received via a user interface, the processor recognizes the user' s speech on the basis of the learned DNN structured parameter by setting a sound-model state set corresponding to the characteristic information if the user' s speech as an output node.
패밀리 해외특허현황	-
유사 선행기술 존재 여부 및 차이점	※ 선행기술 - 특허명: Restructuring deep neural network acoustic models (미국 마이크로소프트, 등록번호: US201403721112A1) - 차이점: 서로 다른 다중 집합 음향적 통계 특성을 수용하는 공통의 DNN 구조 파라미터와 다중 집합 상태 클러스터를 수용하는 DNN-HMM 구조를 가지면서도, 선택된 특정 집합의 입력 음성에 대해 더 적은 수의 상태 집합 출력 노드를 대상으로 훈련/인식을 수행함으로써 더 최적화된 DNN-HMM 구조의 음향모델 파라미터를 학습 하고 음성인식에 사용할 수 있는 방법을 제안함

[3] 시장성

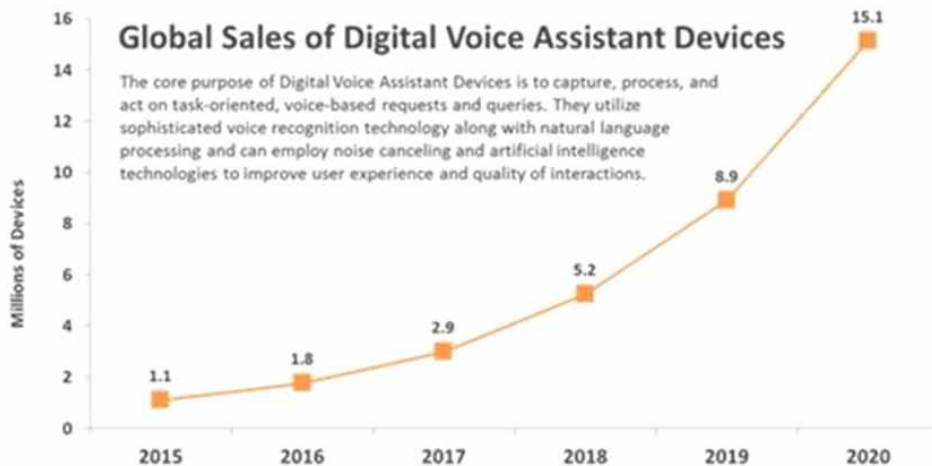
□ 후보기술 관련 시장 개요

- 국내 음성대화처리 기술연구는 ETRI, 포항공대, 서강대, KT등을 중심으로 활발히 진행되고 있으며, 휴대폰 음성인식, IPTV 음성인식검색시스템, 내비게이션등 음성인식시스템에 적용되어 사용됨
- 국외의 경우, Nuance에서 Apple Siri에 핵심인식기술을 공급하여 Google, Microsoft, Apple등을 비롯해 세계 전화자동 응답 장치(ARS)시장의 97%를 점유함



<음성인식 기술의 현재와 미래>

- 음성 인식기술이 적용된 디지털 음성 비서 기기의 경우 시장조사업체인 스트래티지애널리틱스(SA)에서 2016년 180만개에서 2020년에는 1,510만개로 출하량 증가를 예상하였고, 디지털 음성 비서 기기가 매우 빨리 성장하는 분야가 될 것이라며 예측함



<디지털 음성지원 기기 판매 성장률>

[출처 : 스트래티지애널리틱스(SA)]

[4] 사업화 전략

□ 후보기술의 지원 필요성

- 최근 스마트 단말 이용이 대중화됨에 따라, 기기연결을 통한 정보를 소통하는 IoT산업에 대한 기대가 커지고 있으며, 사물인터넷과 인간이 소통하는 Internet of Things to Human(IoT2H)기술의 핵심 원천기술인 자유 발화형 음성대화 처리기술을 확보하는 방향으로 연구가 가속화 되고 있음
- 음성언어기술은 인간의 가장 중요한 정보전달 및 의사소통 수단인 언어를 다루는 기술로서, 21세기 정보화 사회의 핵심 User Interface/User eXperience(UI/UX)기술이며, IT교육 융복합 기술 및 서비스 확보를 통해 음성언어 기술기반 SW는 국가 및 기업의 미래 경쟁력 향상에 기여할 것으로 기대함
- 응용 분야 중 영어 교육 분야의 비즈니스 모델을 통하여 사업화 하는 경우 영어발음 교정 및 음성인식 기술을 교육 콘텐츠 제작업체에 기술이전하고, 학교나 학원 등 대형 오프라인 교육 업체의 커리큘럼에 맞춰 커스터마이징함으로써 기존 학원 고객 리스트를 확보하고 모바일 교육 등으로 다각화 하여 시장 확대하는 사업화 전략을 고려할 수 있음



<음성언어 기반 SW의 응용분야>

- 자유 발화형 음성대화처리 기술은 HCI의 핵심원천 기술로서, 가상 개인 도우미, 지능로봇, 스마트 조연자, 대화형 정보서비스 등 유망 기술 및 서비스 분야에서 폭넓게 사용됨으로써 경제적 파급 효과가 클 것으로 예상