



2018 정보 접근성 기술 컨퍼런스



## “언어지능과 접근성” - ITU-T 표준화 현황

2018. 11. 30.

한국전자통신연구원

언어지능연구그룹 최미란



# 목 차

---

**1. 언어지능과 접근성**

**2. MU-T 언어지능 표준**

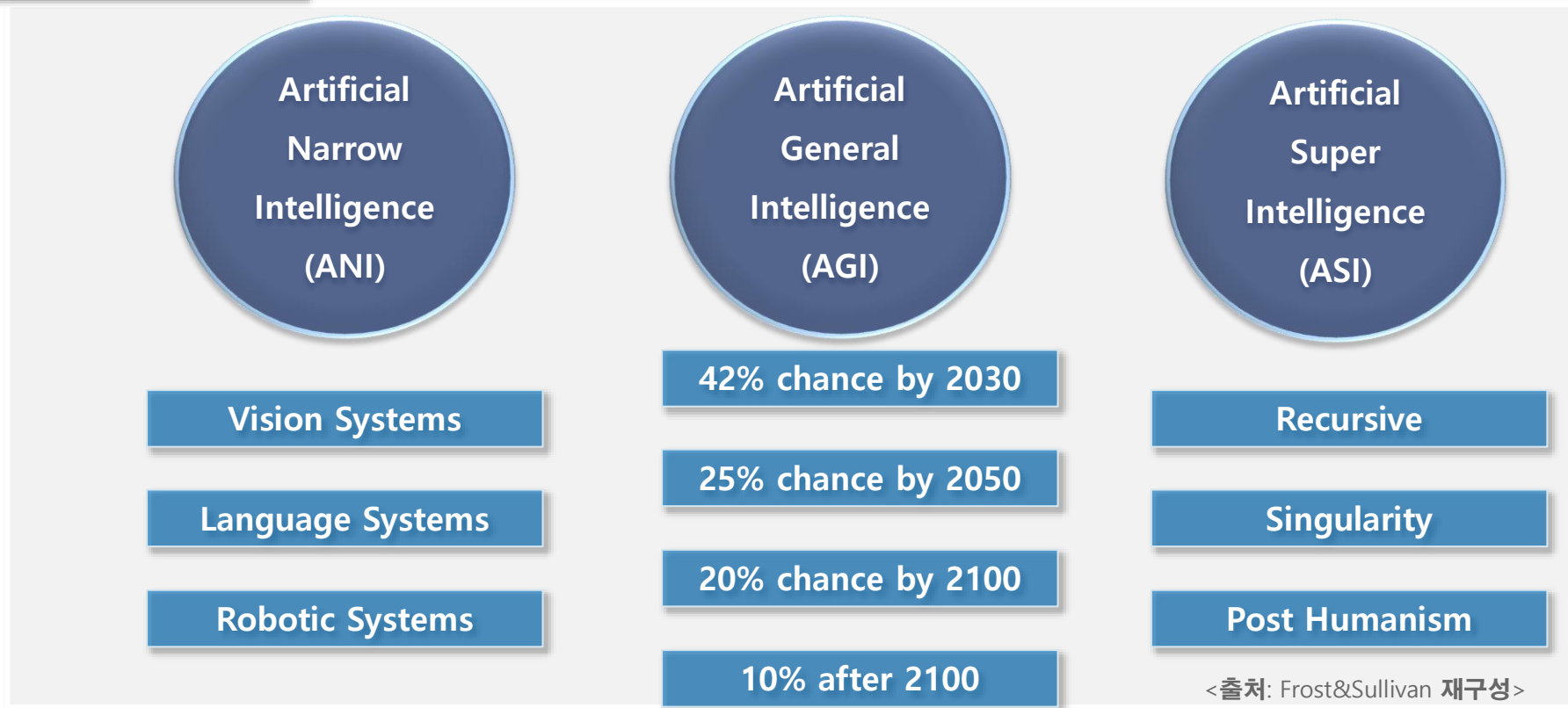
**3. MU-T 접근성 표준**

**4. 맺음말**



# AI 기술 현황

## 3 Types of AI



▪ ANI는 보통 SW로 지칭됨. 현재는 ANI 수준

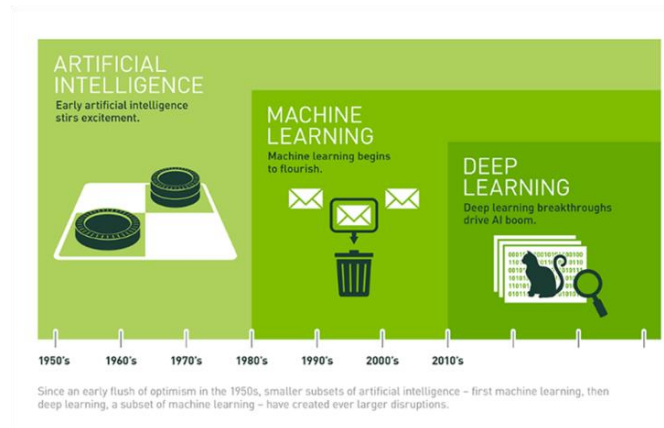
▪ AGI까지는 먼 상황이며 로드맵을 작성하기엔 알 수 없는 미지수들이 많음



## 현재 **혁신적 인공지능 기술 발전**을 주도

[인간이 사전에 고정한 모델에 의존 없이 **학습을 통해 스스로 진화**]

- 음성과 이미지 인식 분야의 획기적 성능 향상
  - 얼굴인식률: 구글 '페이스넷(FaceNet)' 99.96%
- **가능성**
  - 추상적 개념 기반 작업의 해결가능성 제시
- **한계**
  - 사물의 이해 측면에서 아직 초보적 수준
- **향후 연구방향**
  - 추론 과정에 대한 체계적 설명 (Explainable AI)
  - 인간 지식에 의존하지 않는 학습 (Unsupervised Learning)



[출처: NVIDIA]



## 언어지능은 접근성을 향상시키는 중점 기술

- 음성인터페이스
  - 음성인식, 음성합성을 통한 접근성 개선
- 질의응답
  - 음성과 텍스트를 통한 지식 질문과 응답
- 자동통역
  - 언어 장벽의 해소의 의미로 접근성 개선 기술
- 감성인식
  - 감성의 인식과 감성 생성으로 자연스러운 UI 제공
  - 감성을 제공하는 아바타를 이용하여 수화의 이해도를 개선: 표정 생성

# 음성인터페이스(음성인식)

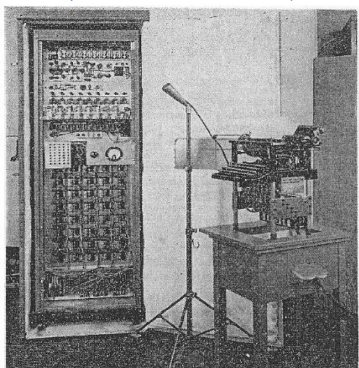




# 음성인식::역사

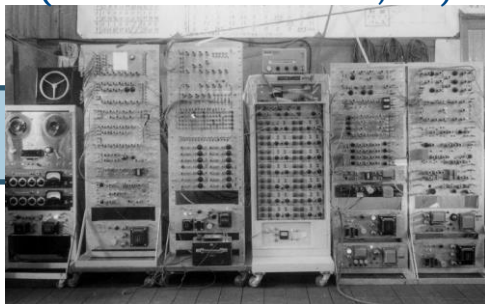
Phonetic typewriter

(RCA Labs, '56)



Spoken digit recognizer

(Radio Research Labs, '61)



모바일 음성검색('10) Siri (iPhone 4S, '11)



Google Home  
Amazon Echo  
MS Cortana  
...

해외



Star Wars (1977)



ARS



Dragon dictation



Google Voice Search



스마트TV

(애플, 구글)



구글 Home, 애플 CarPlay



아마존 Echo



Siri (iPhone 4S)



1970

1980 ... 1990

2000

2010

2012



삼성 애니콜



증권정보  
ARS



음성인식  
내비게이션



모바일  
음성검색



고객센터녹취  
자동통역



스마트TV 스마트스피커  
(삼성, LG) (SKT, KT, ...)

국내

# 음성인식::해외시장

Google, Nuance 등 소수 기업이 전세계 시장 과독점

- Google은 검색, 자동통역 서비스 등 음성언어 기술을 이용한 지식 서비스 사업 확대
- Apple, Google, Amazon 등은 음성대화 기반의 AI 서비스 상용화 경쟁
- 정부 차원의 조직 운영 및 프로젝트 수행 (미국의 LDC, 유럽의 METANET4U 등)

## 주요 핵심 시장

- AI 비서 (assistant) : Google Assistant, Apple Siri, Amazon Alexa, ...
- 인포테인먼트 (infotainment) : Apple CarPlay, Google AndroidAuto, Nuance Dragon drive, MS Windows in the Car, iFLYTEK Auto Speech System, Baidu CarLife, ...
- 음성분석 (speech analytics) : Nuance, Verint, NICE, Genesys, ...



AI 비서



인포테인먼트



음성분석

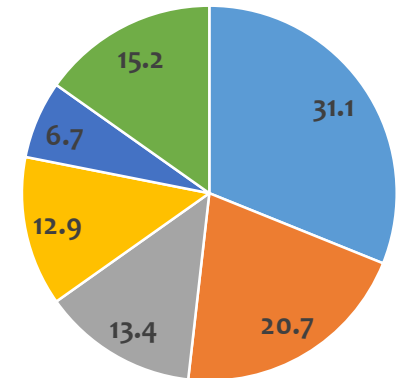
### Nuance

- 구글 대비 열세인 Cloud기반 서비스를 Apple 및 IBM과 협력하여 추진

### Google

- 실사용자 빅데이터 기반으로 음성언어 서비스의 강자 지위
- 100+개 언어 자동번역
- 70+개 언어 음성검색
- 다양한 지식 서비스로 확장

ASR Market Share



■ Nuance ■ Google ■ Microsoft ■ Apple ■ IFLYTEK ■ 기타

[ReportsnReports, '16]



## ETRI 등 국책연구기관에서 중장기적 연구

- 음성검색 서비스를 국내 음성인식 기술을 이용하여 상용화
- 한국어 기반 자동통역 성능은 구글 수준을 상회
- 삼성, LG, SKT, 카카오, 네이버 등에서 내재화된 음성인식 엔진 보유

한국어는 외국대비 경쟁 우위, **다국어 기술은 다국적 기업에 열세**

- 세계 시장은 물론 국내 시장 경쟁력 저하 요인임
- 수출형 제품인 스마트폰, TV, 자동차 등의 UI/UX를 위하여 다국어 솔루션 필요

**국내 시장은 세계시장 대비 협소**하여 관련 기업이 영세

- 악순환 구조 : 시장 협소 → 기업 영세 → 기술 투자 소극적 → 다국어 솔루션 미확보 → 국내 시장에만 주력
- 국내 기업 스스로는 미국과 같은 주도적 기업으로의 성장이 어려움



## 인공지능과 음성인식 기술을 견인하는 개방형 생태계 주도



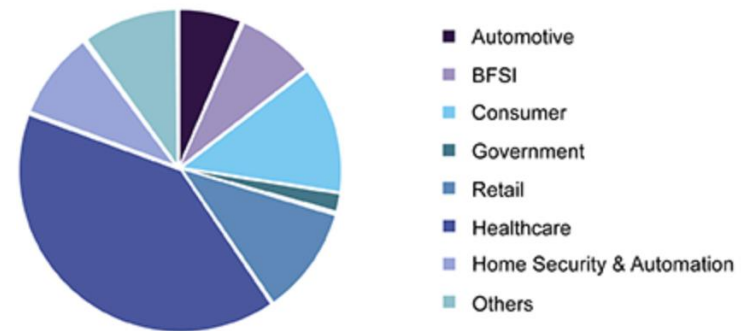
구분	기관	서비스	내용
Open API	구글	<ul style="list-style-type: none"><li>Android API: Android.speech.RecognitionListener</li><li><a href="https://cloud.google.com/speech-to-text/">https://cloud.google.com/speech-to-text/</a> (60분 무료, 15초당 \$0.006, 동영상은 \$0.012)</li></ul>	
	아마존	<ul style="list-style-type: none"><li>AVS: Echo 기반 customized service</li><li><a href="https://developer.amazon.com/alexa-voice-service">https://developer.amazon.com/alexa-voice-service</a></li></ul>	
	카카오	<ul style="list-style-type: none"><li>Newtone API: TTS, STT 지원</li><li><a href="https://developers.kakao.com/features/platform#음성">https://developers.kakao.com/features/platform#음성</a></li></ul>	
	네이버	<ul style="list-style-type: none"><li>Open API: STT, TTS 지원</li><li><a href="https://developers.naver.com/docs/clova/api/#/CSR/API_Guide.md#clova-speech-recognition-api-가이드">https://developers.naver.com/docs/clova/api/#/CSR/API_Guide.md#clova-speech-recognition-api-가이드</a></li></ul>	
	ETRI	<ul style="list-style-type: none"><li>공공 AI open API/Data 포털 : STT, 엑소브레인 등</li><li><a href="http://aiopen.etri.re.kr">http://aiopen.etri.re.kr</a></li></ul>	
	Microsoft	<ul style="list-style-type: none"><li>Cortana</li><li><a href="https://developer.microsoft.com/en-us/windows/projects/campaigns/cortana-skills-kit">https://developer.microsoft.com/en-us/windows/projects/campaigns/cortana-skills-kit</a></li></ul>	
Open source	Johns Hopkins Univ.	<ul style="list-style-type: none"><li>Kaldi : <a href="http://kaldi.sourceforge.net/about.html">http://kaldi.sourceforge.net/about.html</a></li></ul>	
	Cambridge Univ.	<ul style="list-style-type: none"><li>HTK : <a href="http://htk.eng.cam.ac.uk">http://htk.eng.cam.ac.uk</a></li></ul>	
	Carnegie Mellon Univ.	<ul style="list-style-type: none"><li>Sphinx: <a href="http://cmusphinx.sourceforge.net/">http://cmusphinx.sourceforge.net/</a></li></ul>	
	{Chatbot}	<ul style="list-style-type: none"><li>MS bit framework, LUIS, IBM Watson, wit.ai, Chatfuel, Manychat, Motion AI, Personify, ChatterBot, ChatScript, BotPress, BotKit, BotMan, ...</li></ul>	

# AI 비서/스피커

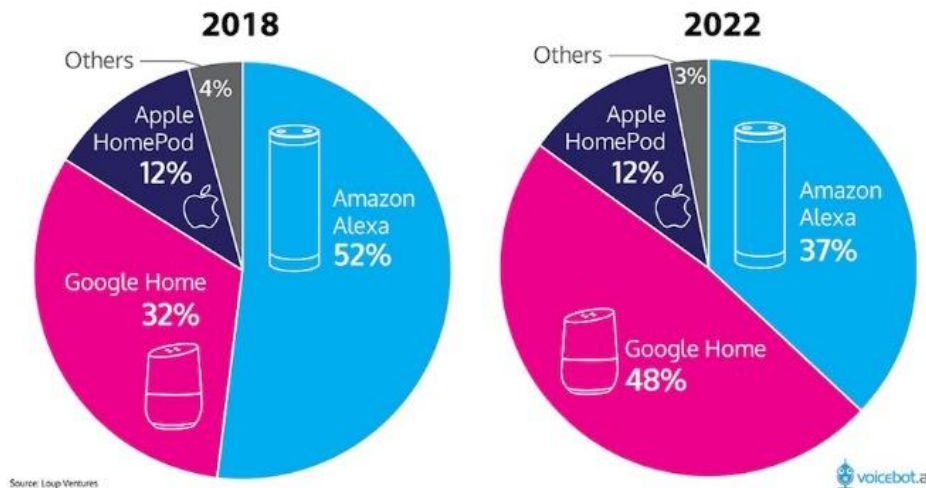
[IDC]

- '20년까지 스마트 운전 산업의 51%, 중국 휴대전화 및 웨어러블 업계의 68%가 대화식 AI 시스템을 채용할 것으로 예측

[Vertical market share, Grand View research, 2016]

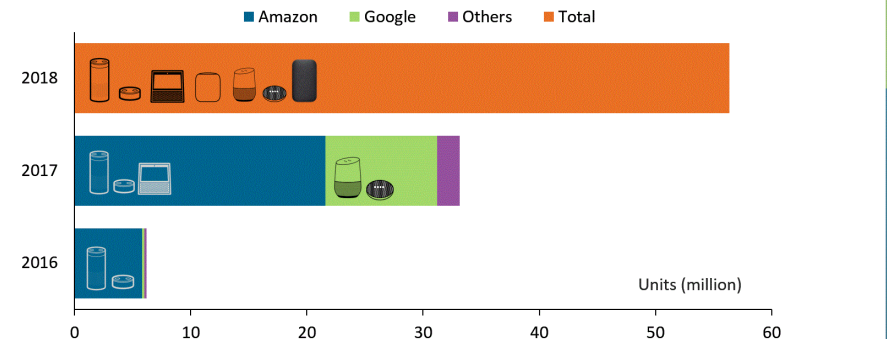


## GLOBAL SMART SPEAKER MARKET SHARE



## 56.3 million smart speakers to ship in 2018

Worldwide smart speaker estimates and forecasts by vendor, 2016 - 2018



Source: Canalis estimates and forecasts, Smart Speaker Analysis, December 2017

# 국외 AI 비서/스피커 현황



(세계시장) '15년 3.6억불 → '20년 21억불 (CAGR 42.3%)



(해외)	아마존	구글	애플	MS	바이두	알리바바
제품	에코 (dot, tap, look, show 등)	홈 (max, mini 등)	홈팟	Invoke	리틀 피시	Tmall Genie
출시일	'14. 11.	'16. 11.	'17. 12.	'17.10.	'17. 4.	'17. 8.
특징	알렉사 Skills 기반 생태계 세계 시장 70% 점유 (780만대, ~'16)	어시스턴트 ('16) 터치스크린 API 공개 ('18.1.8.)	시리 ('11)	코타나 ('14)	DuerOS ('15) 스크린 탑재	AliGenie
서비스	음악, 요리, 택시 호출 E-commerce IoT (home, GE/월풀) Ford SYNC, HMC BlueLink Skill	검색 음성통화 음악 IoT (Nest) 항공/교통정보	검색 (Wolfram Alpha) 음악 팟캐스트	검색 (Bing) 음악 교통 IoT (home) 스카이프	온라인 주문 날씨 음악 IoT (home)	온라인 주문 날씨 음악 IoT (home)
수익 모델	아마존 페이 -쇼핑 (수수료) -구독 (음원/영상)		애플 페이			
언어	2+	8+	20+	8	1	1

# 국내 AI 비서/스피커 현황



(국내시장) '20년 11조원



(국내)	SKT	네이버	카카오	삼성전자	LG전자/U+	KT
제품	누구	웨이브	미니	빅스비	씽큐 허브 U+ 우리집AI	기가지니
출시일	'16. 9.	'17. 7.	'17. 9.	'17. 3.	'17.11.	'17. 1.
특징	월 약 1만대 판매	AI 플랫폼 '클로바' 'Face' 추진중	카카오 플랫폼		네이버 제휴	IPTV STB
서비스	멜론, 스마트홈, 일정, 알람, 날씨, 팟캐스트, 주문배달, 무드등, 라디오, 위키검색, 길안내, 멜론어린이, 11번가, 프로야구, 운세 등	네이버뮤직, 팟캐스트, 번역, 콘텐츠 추천, 검색 등	카카오톡, 멜론, 택시, 내비, 다음포털 서비스 등, 결제	- 일정, 뉴스, 앱 이용, 콘텐츠 추천 - 사물, 이미지, 텍스트, QR 코드 등을 인식해서 정보 제공 (번역, 쇼핑 등)	VOD, 전자제품제어, 검색, 영어교육, 쇼핑 등	배달음식, 지니뮤직, 알람, 뉴스검색, 교통정보, 라디오, 일정, 버스조회, 카카오택시, 추가, 키즈콘텐츠 등

# 음성인식::발전방향

## 음성대화 이해 기술 → 차세대 UI/UX의 핵심 요소



### 기술 확산 전망



#### 음성인식

“다양한 컴퓨팅 환경에서 음성인식과 같은 natural interface 확산” Microsoft, 빌게이츠



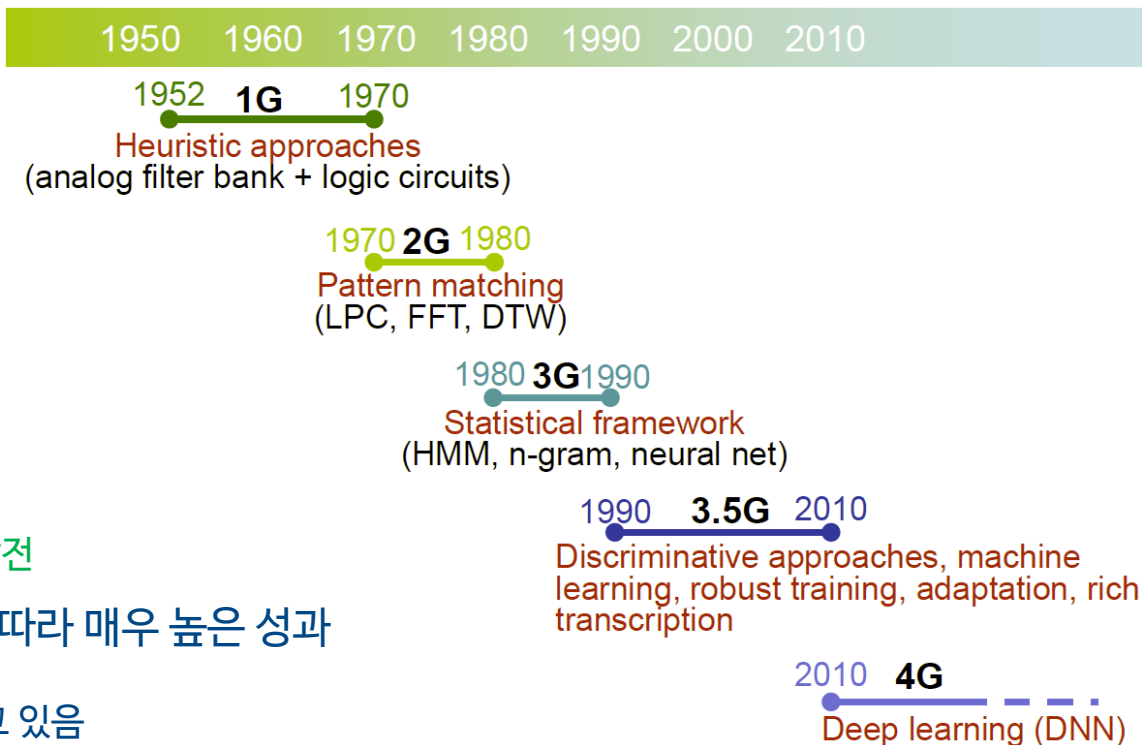
#### 대화처리

“자연어 이해야 말로 인공지능 기술의 핵심” Google, 레이 커즈와일



# 음성인식::현안

<S. Furui, '15>



- (明) 기계학습 및 컴퓨터 기술의 도움으로 비약적 발전
  - 특히 최근 '딥러닝'을 사용하는 DNN에 따라 매우 높은 성과
- (暗) 여전히 해결할 수 없는 많은 도전 과제를 가지고 있음
  - 기계학습 기반 다양한 지식원 구현 → 인간의 예측 방식에 대해 깊이 생각하고 모델링
  - 빅데이터 생성 및 사용 방법론 → 다양한 지식원을 유연하게 활용
  - 인간 수준으로 의미/의도 이해를 모델링하고 처리하는 방법
  - 능동학습, 비교사/반교사 학습 및 적응 기술

# ETRI 음성인식



# 핵심 기술 R&D



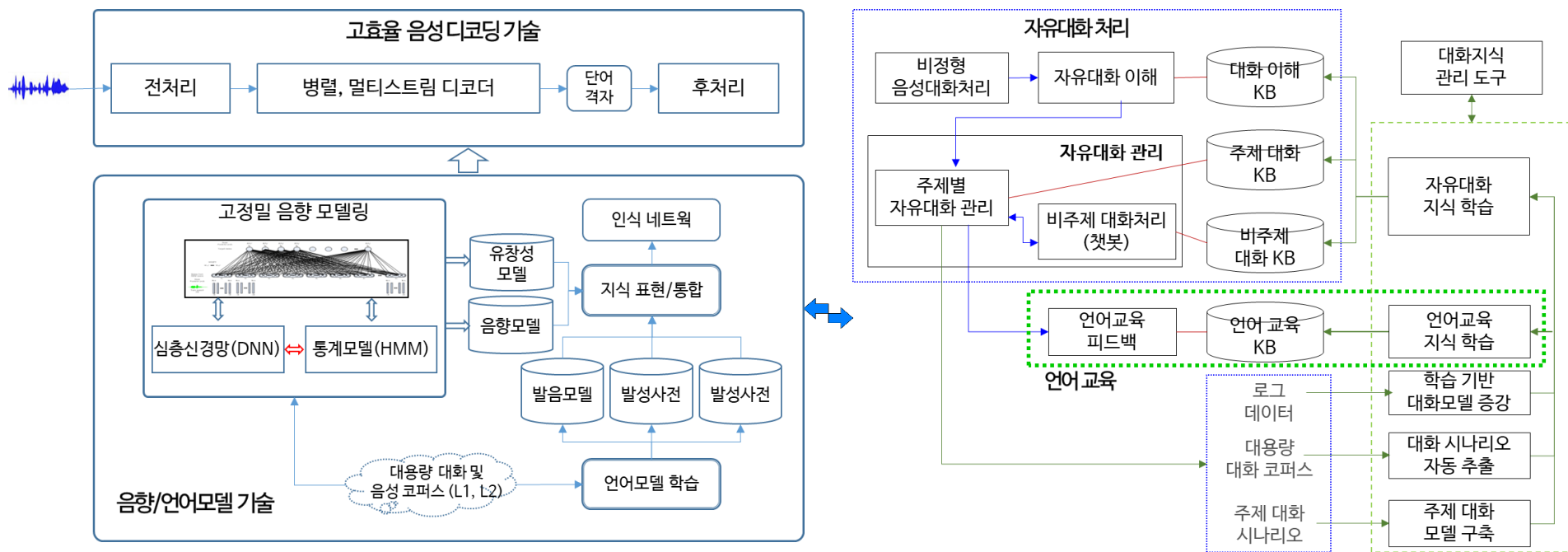
## 음성대화 기술 ← 자연어 { 음성인식 + 대화처리 }



# 음성대화처리 요소기술 개발



## 음성인식 및 대화처리 기술



# 영어 교육



## 영어 말하기 학습 서비스 사업화

- 영어 말하기 학습 솔루션
- SRI, Nuance 등 국내외 솔루션과의 비교 평가를 통해 선정/사업화
  - ◆ 스마트폰용 영어학습 서비스('12)
  - ◆ 영어학습 음성인식 게임 SW('12)
  - ◆ EBS 초목달 스피킹 서비스('15)



〈MWC '16, '17, '18 전시〉



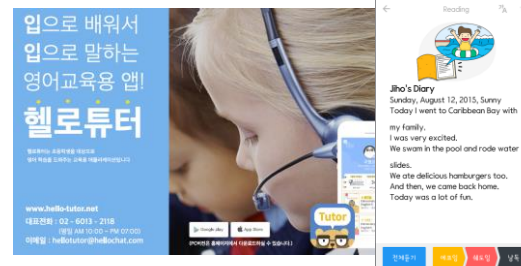
〈NC소프트 - 호두잉글리시〉



〈EBS 초목달〉



〈지앤비교육 - GnB English〉



〈헬로튜터〉



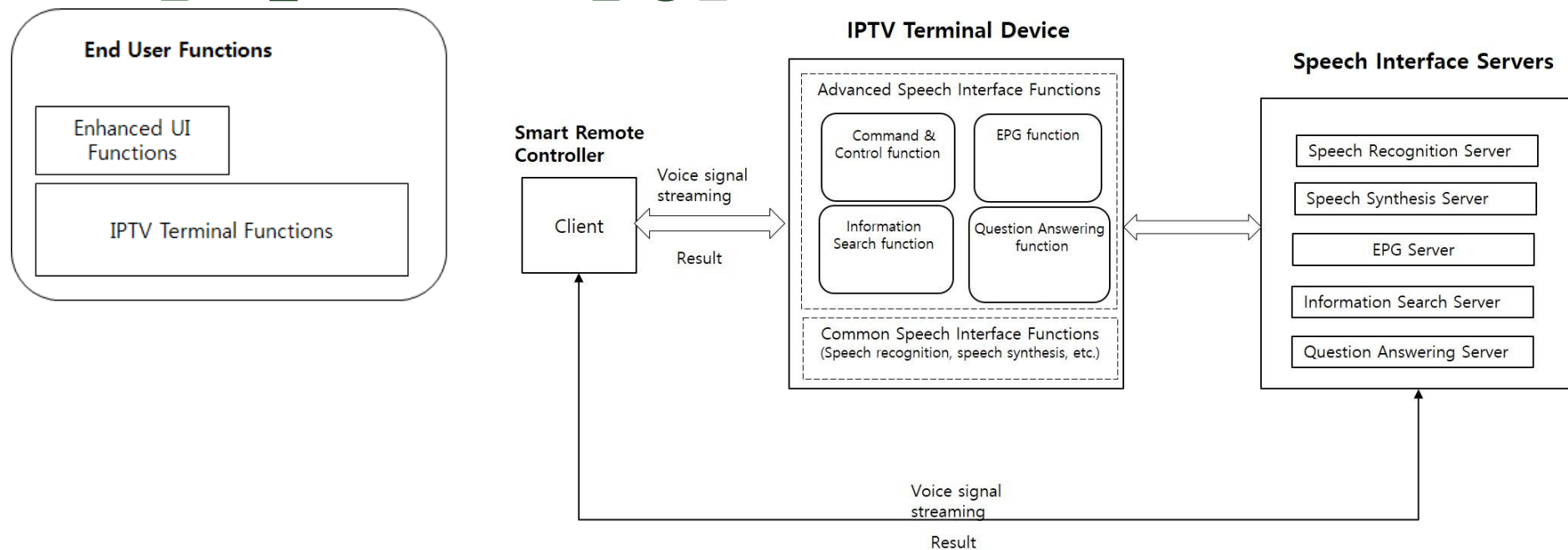
〈잉글리시아이〉



# 음성 인터페이스 표준

## IPTV 음성 인터페이스 표준화

- ITU-T SG16 Q13&26
- H.IPTV-EUIF: Enhanced user interface framework for IPTV service
- 표준화 내용: 고기능 IPTV 사용자 인터페이스를 위한 요구사항 및 기능 정의, 음성인터페이스, 터치 UI, 제스처
- 2016년 11월 H.703으로 표준 승인







# 언어 학습 표준

## 표준화 진행 사항

- 표준화 내용: 음성인터페이스 기술과 자연어처리 기술을 활용한 언어 이러닝 시스템 표준화
- ITU-T SG16 Q21에서 신규 표준항목 승인, 2017년 10월 표준 승인.
- ITU-T F.746.5: Language e-Learning service framework based on speech/NLP technology
- 후속표준: 언어이러닝을 위한 대화 시스템 (ITU-T SG16, 2018.03, 신규 승인)



# 질의응답 (언어이해&엑소브레인)



# 언어 이해 - 엑소브레인

## 언어 인공지능 SW



미국의 대표적 공업지대인 미국 중서부와 북동부 지역을 뜻하며, 미국 제조업의 몰락을 상징적으로 보여주는 말로 사용되어 온 용어입니다. 자동차 산업의 중심지인 디트로이트를 비롯해 철강 산업의 메카 피츠버그, 그 외 필라델피아, 볼티모어, 멤피스 등이 속해 있는 지역으로, 트럼프의 45대 대통령 당선의 일등공신이 된 이곳은 어디일까요?

[장학퀴즈 <대결! 엑소브레인> 문제]



# 언어 이해 - 엑소브레인

## 개념 및 내용

### 전문직 종사자와 지식 소통이 가능한 인공지능 SW 개발

\* 법률/특허 등의 분야에서 질의응답(Question Answering)을 통한 지식 서비스 제공

#### Question Answering

- 금융/민원 등 전문지식 상담

신장결석으로  
충격파쇄석수술을 했는데,  
수술시 보험금 지급이  
가능한가요?



자연어 심층이해   자연어 지식 생산



자연어 질의응답

#### Decision Support

- 법률/특허 등의 전문가 판단 지원

이 특허출원서와 중복되는  
선행특허는?



#### Smart Advisor

- 전문직종 의사결정 지원

Edward Fox was diagnosed  
with type 2 diabetes last  
year. What are good  
medications for him?



#### Thinking Machine

- 지능형 단말의 인공지능으로 활용

착륙예정점에 대한 경사각 수직면  
유도 결함의 정비 항목은?





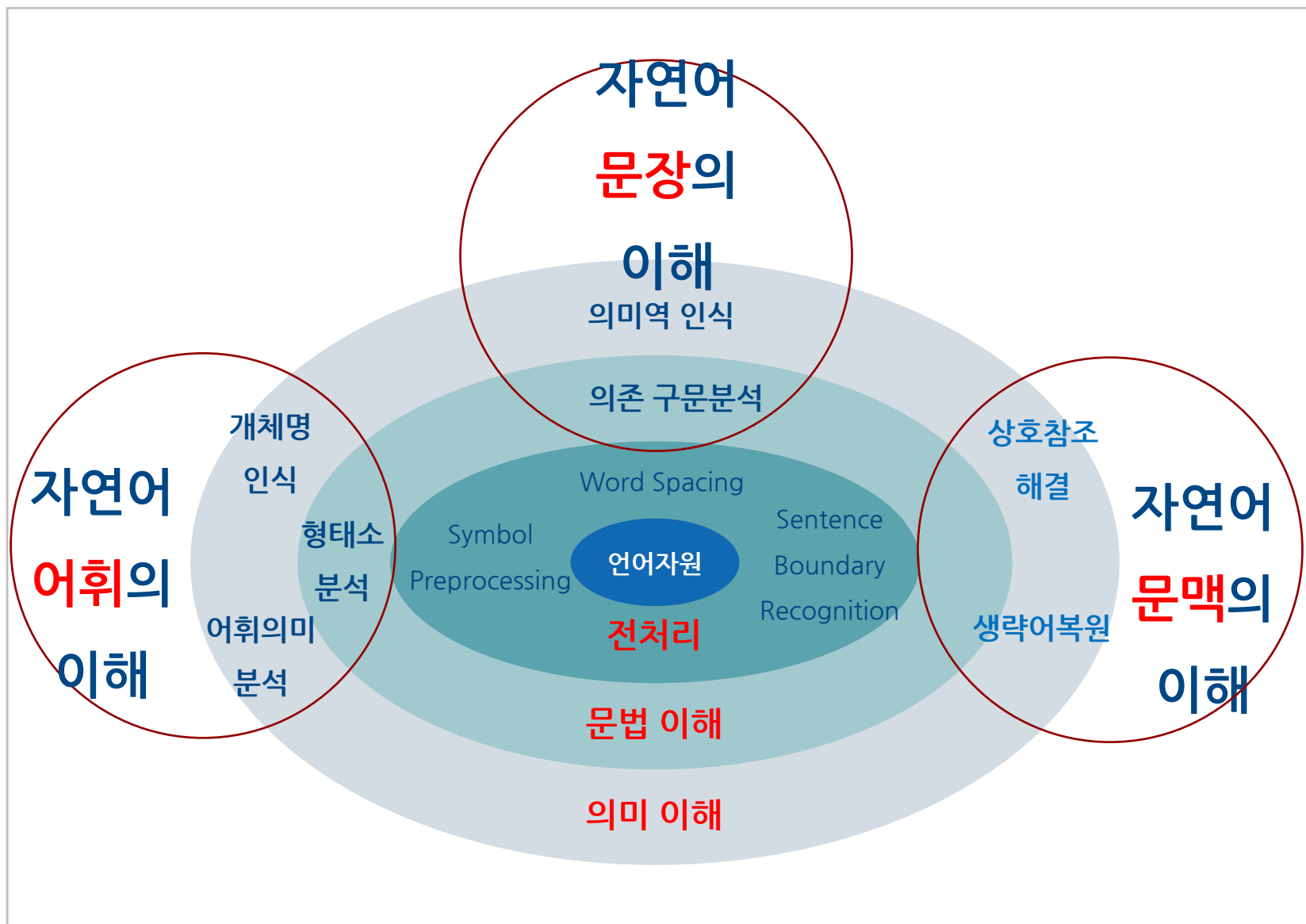
## 핵심 기술

**복잡한 자연어로 기술된 문제의 의미를 이해하고, 정답을 추론하여 생성함**





# 1) WiseNLU 언어이해 기술 개요



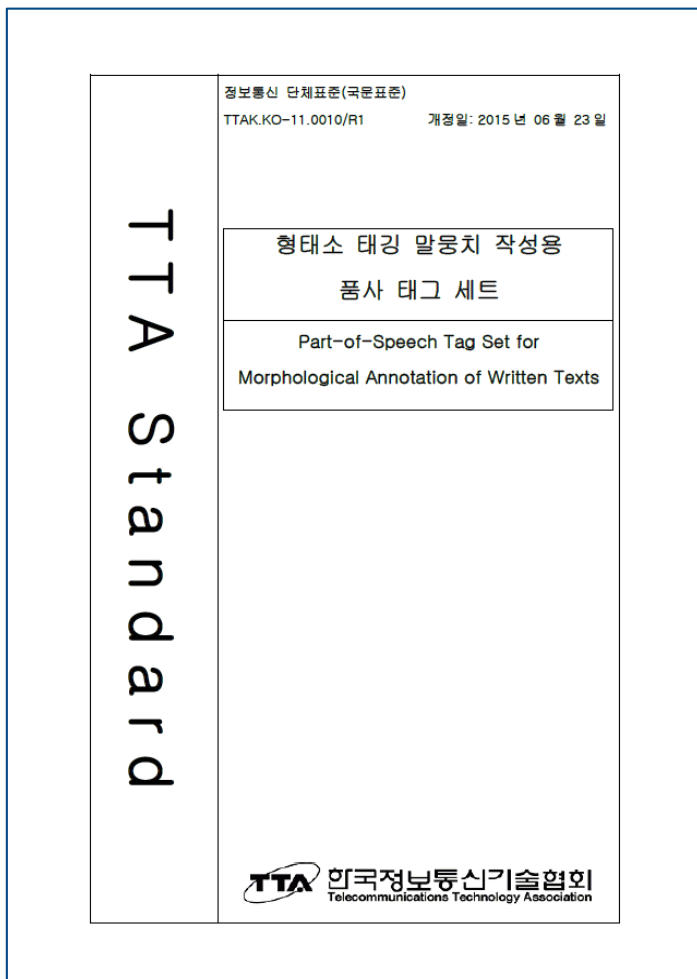


# 문법분석의 의미분석



## 형태소분석

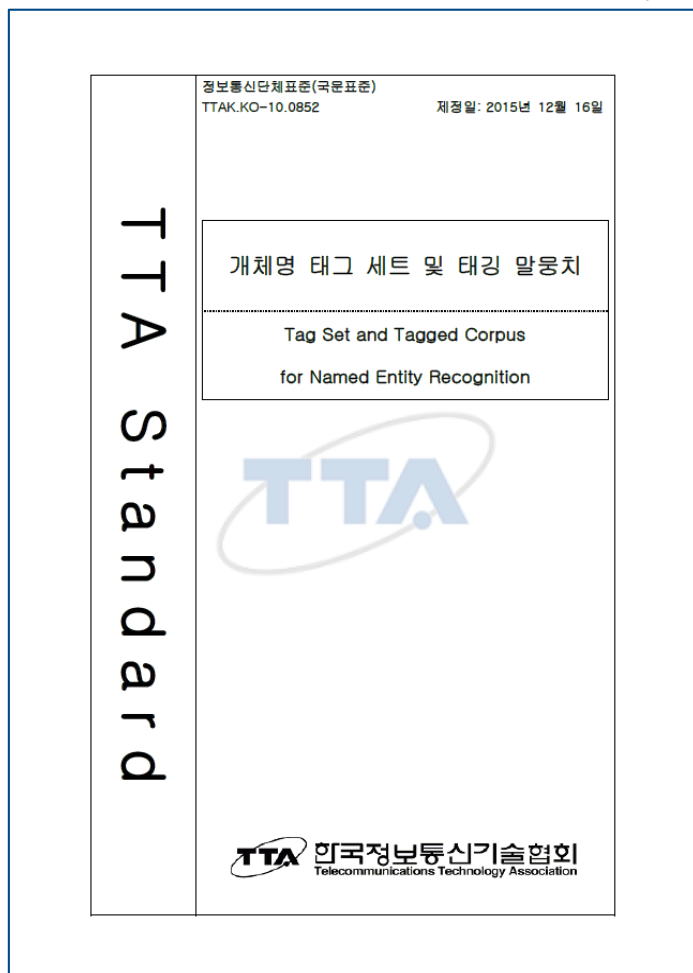
### ■ 정보통신 단체표준 TTAK.KO-11.0010/R1



대분류	소분류	세분류	세종태그
(1) 체언	명사	일반명사	NNG
		고유명사	NNP
		의존명사	NNB
	대명사		NP
(2) 용언	수사		NR
	동사		VV
	형용사		VA
	보조용언		VX
	지정사	긍정지정사	VCP
		부정지정사	VCN
(3) 수식언	관형사		MM
	부사	일반부사	MAG
		접속부사	MAJ
(4) 독립언	감탄사		IC
(5) 관계언	격조사	주격조사	JKS
		보격조사	JKC
		관형격조사	JKG
		목적격조사	JKO
		부사격조사	JKB
		호격조사	JKV
		인용격조사	JKQ
	보조사		JX
	접속조사		JC

## 개체명 인식

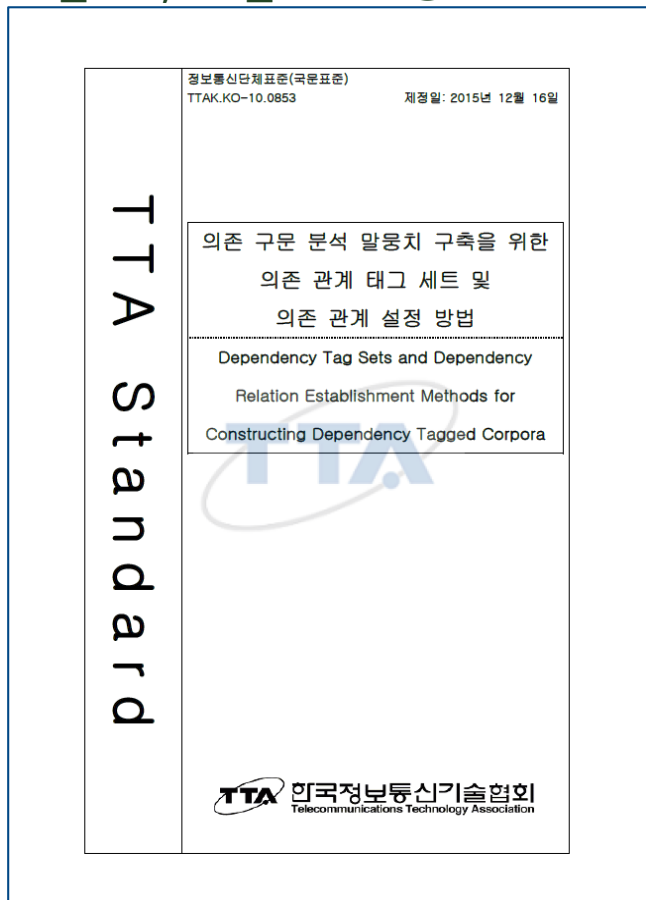
- 정보통신단체표준 TTA.KO-10.0852
  - 총 146개의 세부 분류 개체명 사용 (대분류 15개)



대분류	설명
PERSON (PS)	인물
LOCATION (LC)	지역
ORGANIZATION (OG)	기관
ARTIFACTS (AF)	인공물
DATE (DT)	날짜
TIME (TI)	시간
CIVILIZATION (CV)	문명
ANIMAL (AM)	동물
PLANT (PT)	식물
QUANTITY (QT)	수량
STUDY_FIELD (FD)	학문 분야
THEORY (TR)	이론
EVENT (EV)	이벤트
MATERIAL (MT)	물질
TERM (TM)	용어

## 의존 구문분석

- 정보통신단체표준 TTAK.KO-10.0853
- 의존관계 태그는 구문 태그와 기능 태그를 결합하여 사용
  - 예: NP\_SBJ, VP\_MOD 등



구문 태그	의미
NP	체언 (명사, 대명사, 수사)
VP	용언 (동사, 형용사, 보조용언)
AP	부사구
VNP	긍정 지정사구 (명사+이다)
DP	관형사구
IP	감탄사구 (호칭 및 대답 등의 표현)
X	의사 구 (pseudo phrase, 조사 단독 어절 또는 기호 등)
L	부호 (왼쪽 괄호 및 따옴표)
R	부호 (오른쪽 괄호 및 따옴표)

기능 태그	의미
SBJ	주어
OBJ	목적어
MOD	관형어 (체언 수식어)
AJT	부사어 (용언 수식어)
CMP	보어
CNJ	접속어 (~와)

## 의미역 인식 국가 표준

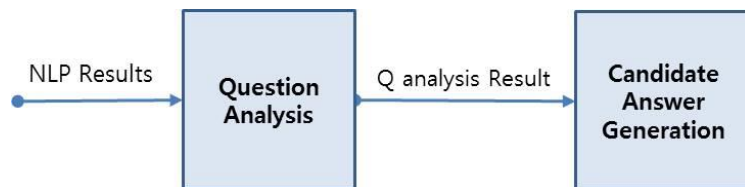
### ■ KS\_X\_NEW\_2016\_1938 (진행 중)

대분류	세분류	정의
ARG-N (필수역)	ARG0	- 서술어의 동작주, 행위자 (agent)
	ARG1	- 서술어의 피동작주, 대상 (patien, theme)
	ARG2	- 시작점, 수익자, 속성 등(starting point, benefactive, attribute)
	ARG3	- 도착점(ending point)
ARG-M (부가역)	ARGM-LOC	- 장소(locatives)
	ARGM-DIR	- 방향(directional)
	ARGM-CND	- 조건(condition)
	ARGM-MNR	- 방법(manner)
	ARGM-TMP	- 시간(temporal)
	ARGM-EXT	- 범위(extent)
	ARGM-PRD	- 보조서술(secondary predication)
	ARGM-PRP	- 목적(purpose clauses)
	ARGM-CAU	- 발생 이유 (cause clauses)
	ARGM-DIS	- 담화 연결 (discourse)
	ARGM-NEG	- 부정 (negation)
	ARGM-INS	- 도구 (instrument)



# ITU-T 질의응답 메타데이터 표준

- F.746.7(구 F.IQAS-META): Metadata for Intelligent Question Answering Service”
  - ITU-T SG16 Q21 에서 진행되었으며 2018년 8월 표준승인 완료
  - ITU-T F.746.3 (2015.11)에서 정의한 질의응답서비스를 위한 메타데이터 정의
  - 4개의 모듈에 대한 메타데이터를 다룸
  - 구성
    - QA 시스템 개요
    - 기능별 구성요소와 정보 흐름
    - 모듈별 메타데이터
      - information for natural language processing
      - information for question analysis
      - information for candidate answer generation
      - information for answer inference/generation



[Table 4] NLP information elements/attributes – Named Entity Analysis

Element / Attribute	Definition	Type	Level
NE	Result of Named Entity analysis	object	
id	Id of a word (starting from 0 in a sentence)	integer	
text	Text of a word	string	
type	Named Entity type Ex. River, person, organization	string	
begin	Id of the first POS tag in word	integer	
end	Id of the last POS tag in a word	integer	
weight	Reliability of WSD analysis from 0 to 1	double	
Common noun	Specify if it is common noun (1) or proper noun (0)	0 or 1	

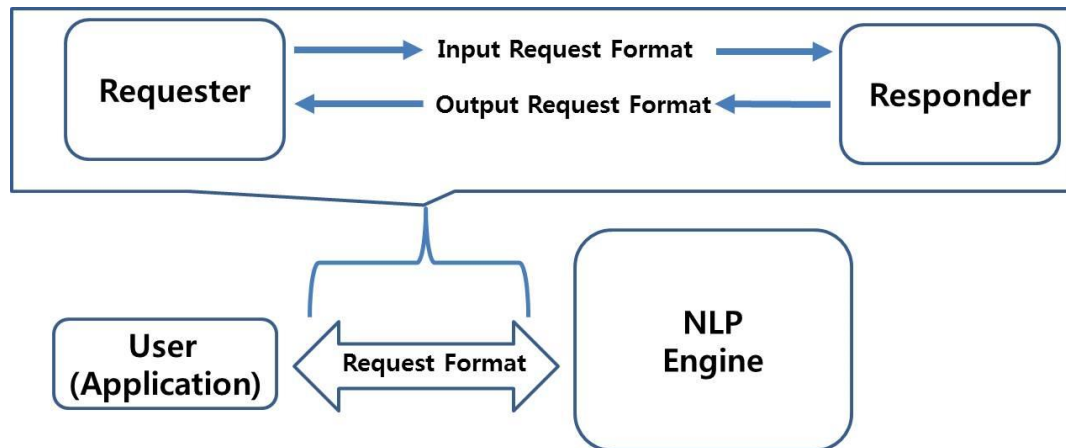




# 자연어처리 표준 - ISO

## NLP Request Format

- 표준화 내용: AI 응용시스템에서 공통적으로 필요한 자연어처리 요청 포맷
- ISO TC37 SC4 언어자원 표준화 그룹에서 프로젝트 리더로 표준화 진행 중
- 기존의 자연어처리 표준과 연결하여 모듈별 상세 내용 정의



### Contents

Page

1	Scope .....	5
2	Normative references .....	5
3	Terms, definitions, abbreviated terms and conventions .....	5
4	Conventions .....	5
5	Structures and Data Model .....	6
5.1	Structure .....	6
5.2	Data Model .....	6
6	Example Scenario .....	7
7	Root Element .....	7
7.1	Introduction .....	7
7.2	Syntax .....	7
7.3	Semantics .....	7
7.4	Example .....	7
8	Data Types .....	7
8.1	SourceDataLocatorType .....	7
8.1.1	Introduction .....	7
8.1.2	Syntax .....	7
8.1.3	Semantics .....	7
8.1.4	Example .....	7
9	Input Format .....	8
9.1	Introduction .....	8
9.2	Syntax .....	8
9.3	Semantics .....	8
9.4	Example .....	8
10	Output Description .....	8
10.1	Introduction .....	8
10.2	Syntax .....	8
10.3	Semantics .....	8
10.4	Example .....	8
11	NLP Type .....	8
11.1	Introduction .....	8
11.2	Syntax of NLP Type .....	8
11.3	Semantics of NLP Type .....	8
11.4	NLP request for Morpho-syntactic annotation framework (MAF) .....	8
11.5	NLP request for Linguistic annotation framework (LAF) .....	8
11.6	NLP request for Lexical markup framework (LMF) .....	8
11.7	NLP request for Syntactic annotation framework (SynAF) .....	8
11.8	NLP request for Semantic annotation framework (SemAF) .....	8
11.9	NLP request for Emotion annotation framework (EmoAF??) .....	8
12	Output Format .....	9
12.1	Output .....	9
12.2	ResultItem .....	9
12.3	SystemMessageType .....	9
12.4	Example .....	9

# 자동통역

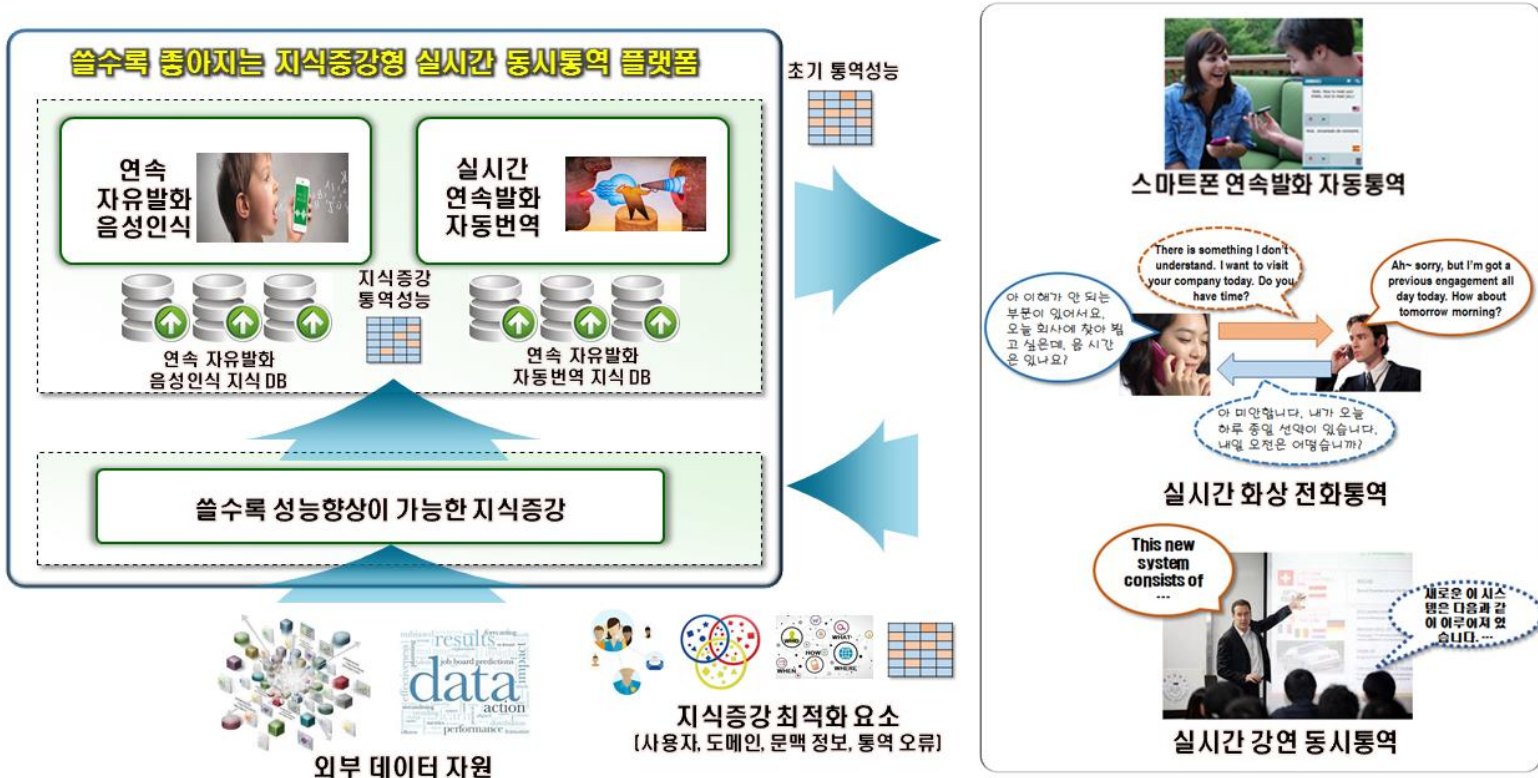




# 실시간 동시통역 - 지니톡

## 개념 및 내용

관광/국제행사/전화/강연/회의/동영상 등에서 연속된 자유발화에 대해 의사소통 단위로 실시간 통역이 가능하고, 쓰면 쓸수록 성능이 개선되는 지식증강형 실시간 동시통역 원천기술 개발





# 실시간 동시통역 - 지니톡

## 핵심 기술

## 기존 자동통역 기술과의 차별점

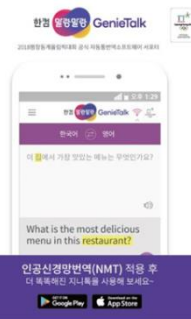
	단문형 자동통역	연속발화 실시간 동시통역
통역 방식	한 문장씩 순차통역	연속 자유발화 동시통역
발성의 자유도	정형 문장 제한발화	비문[非文]을 포함한 자유발화
번역지식 구축 방법	수작업으로 구축된 번역지식 /패턴을 추가 적용하여 개선	번역지식 자동추출 및 증강으로 번역 성능의 점진적 개선 가능함
어휘 표현력	10~30만 단어급	50만 단어 이상(사실상 무제한 어휘)
적용 도메인	여행 분야	일상 분야
서비스방식	스마트폰 기반	강연 통역, 동영상 통역, 화상 통역 등
주요 기술	<ul style="list-style-type: none"> <li>- 단일 발화문 음향/언어 모델링 기술</li> <li>- 규칙/통계/신경망 기반 자동번역 기술</li> </ul>	<ul style="list-style-type: none"> <li>- 딥러닝 기반 연속 자유발화 음향/언어 모델링 기술</li> <li>- 비정형 언어분석 및 의사소통 단위 실시간 분절 기술</li> <li>- 신경망 기반의 다중모델 하이브리드 자동번역 기술</li> <li>- 언어 및 도메인 확장을 위한 Multi source learning 및 transfer learning 기술</li> <li>- 자유로운 양방향 대화가 가능한 Zero UI 통역 기술</li> </ul>



# 실시간 동시통역 - 지니톡

## ●한컴인터프리 말랑말랑지니톡 상용서비스 실시(2016.7)

- ETRI 지니톡 대국민 서비스(2012~2015)를 통해 220만명 다운로드 기록
- 한/영, 한/중, 한/일, 한/스, 한/불 상용 서비스 실시 중, 한/독, 한/러, 한/아랍어 포함하여 **평창동계올림픽 8개국어 통역 서비스 지원**



## ●신경망 기반 자동번역(Neural Machine Translation) 기술 구현 및 서비스 적용(2017.2)

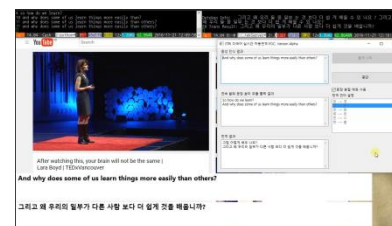
## ●웨어러블 통역 응용 기술 개발

- 벅밴드 이어셋 웨어러블 자동통역 기술 선보임(2017 MWC)
- ETRI **“제로 유아이(Zero UI) 자동통역 기술” ISO 국제표준 승인(2017.8)**



## ●동영상 강연 통역, Live 강연 동시통역 프로토타입 시스템 구현

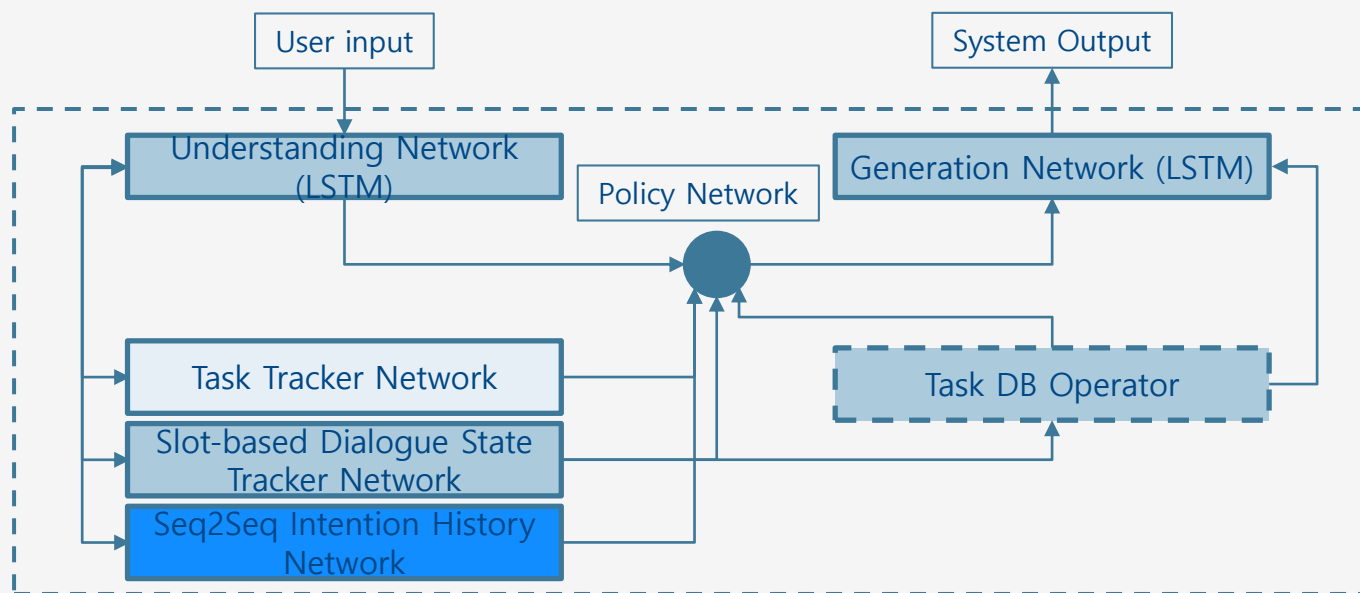
- 실시간 동시통역 서비스(강연, 동영상, 화상전화 동시통역 등) 발굴 진행 중





# 통역용 대화 관리 기술

## ● DNN 기반 대화시스템의 예



# 통역용 DNN 기반 사용자 의도 인식

## ● DNN 기반 모델 테스트 및 엔진 적용

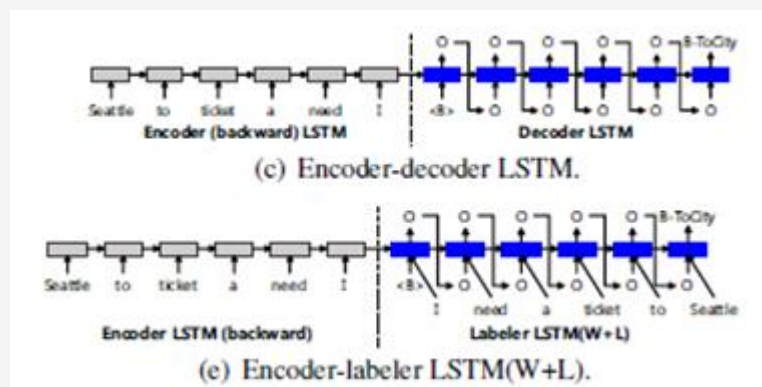
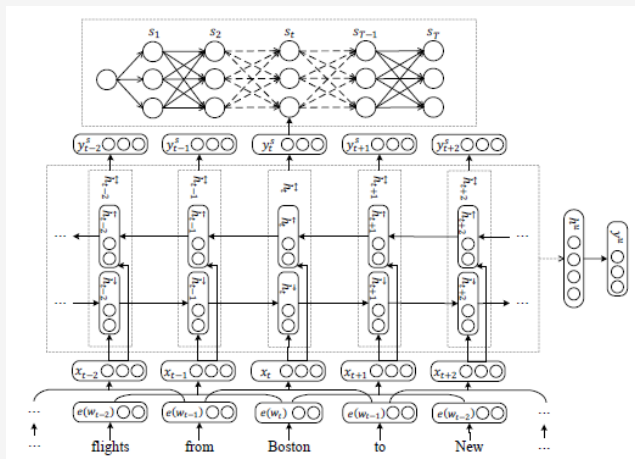
### - DNN 기반 slot tagging 기술 (NER, 비정형처리 등 공통 적용)

#### ■ seq2seq 모델

- Encoder-labeler LSTM
- focus mechanism
- NMT unknown 처리 응용 (copy mechanism 등)

### - Joint 모델

#### ■ RNN, Attention-based RNN



Model	ATIS		CQUD	
	Intent	Slot	Intent	Slot
ID only	97.53	—	95.34	—
SF only	—	95.14	—	85.78
Pipeline	97.53	95.41	95.34	86.96
Joint (equal)	98.10	95.49	96.05	87.12
Joint (ID oriented)	98.10	95.49	<b>96.35</b>	86.63
Joint (SF oriented)	97.87	<b>95.61</b>	95.93	<b>87.23</b>

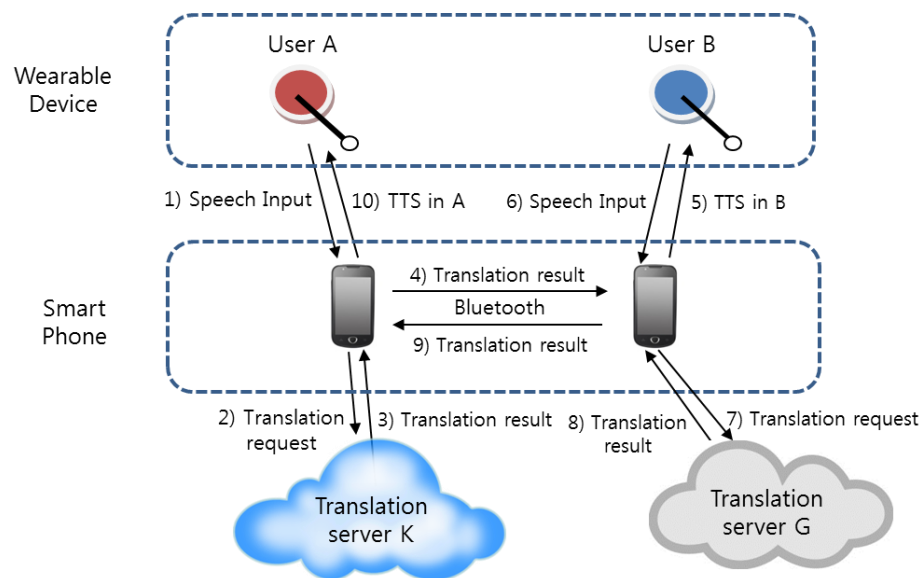
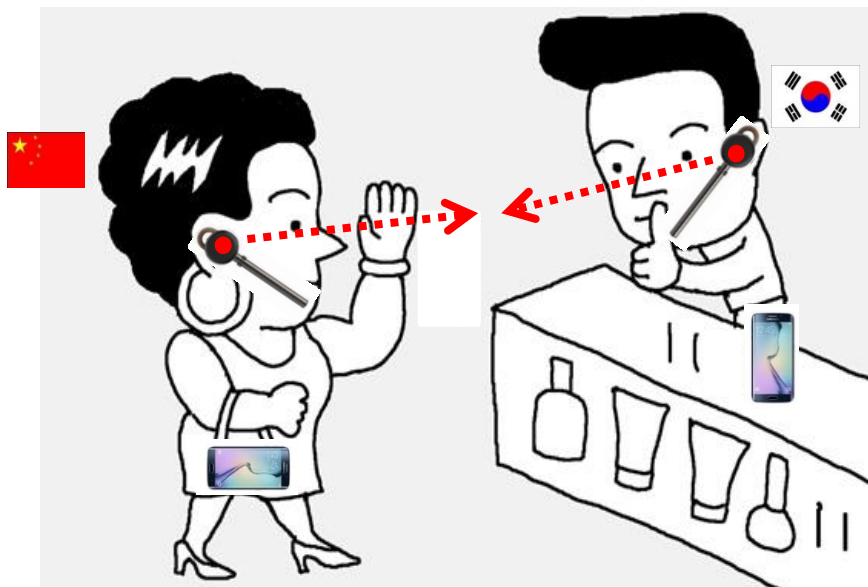
Table 3: Comparison of joint model and separate model



# 대면상황 음성통역 표준

## 표준화 진행 사항

- 표준내용: 대면상황 음성통역시스템의 User Interface와 시스템 표준
- ISO 20382-1 and 2: Face-to-face speech translation
- ISO JTC1 SC35 User Interface 표준화 회의에서 표준 승인 (2017.9)
- 후속 표준으로 동시통역시스템 표준화 진행 중
- ITU-T SG16 Q24 Human Factors 그룹에서 통역평가표준 (E.FAST) 진행 중







# 언어지능 기술 로드맵

2017

2018

2019

2020

2021

2022

엑소브레인  
(언어)



전문분야 QA 기술



한국어/영어 QA 기술

지니튜터  
(음성)



대화형 언어학습 기술



자유대화형 정보서비스(상담, 교육, 오락)



인공지능 상담사

지니톡  
(음성/언어)



평창동계올림픽 8개국 자동통역 기술



실시간 동시통역 기술



실시간 방송 통역 기술

# ITU-T 인적요소 및 접근성 표준화





# ITU-T SG16 Q.24 인적요소 (Human Factors) 그룹

- Title: Human factors related issues for improvement of the quality of life through international telecommunications
- Motivation:
  - Better understanding of human factors that will accommodate persons with specific needs
  - Improve usability through ICT
- Study items:
  - The basic elements of the user interface that are relevant to multiple services
  - Methods for a smooth dialogue between the user and the service
  - Approaches to facilitate entering alphabetic information (keyboards)
  - Language-specific issues
  - symbols, pictograms and emoticons to be available to telecommunication/ICT users
  - issues of human factors for the new technologies



# 인적 요소 (Human Factors)의 정의

- Human factors examines **the relationship between human beings and the systems** with which they interact by focusing on improving efficiency, creativity, productivity and job satisfaction, with the goal of minimizing errors. [WHO]
- Ergonomics (or human factors) is the scientific discipline concerned with the **understanding of interactions among humans and other elements of a system**, and the profession that applies theory, principles, data and methods to design to optimize human well-being and overall system performance. [International Ergonomics Association]
- **Human factors and ergonomics** (commonly referred to as **human factors**) is the application of psychological and physiological principles to the (engineering and) design of products, processes, and systems. The goal of human factors is to reduce human error, increase productivity, and enhance safety and comfort with a specific focus on the interaction between the human and the thing of interest [Wikipedia]



# 인적 요소 그룹의 표준화 항목

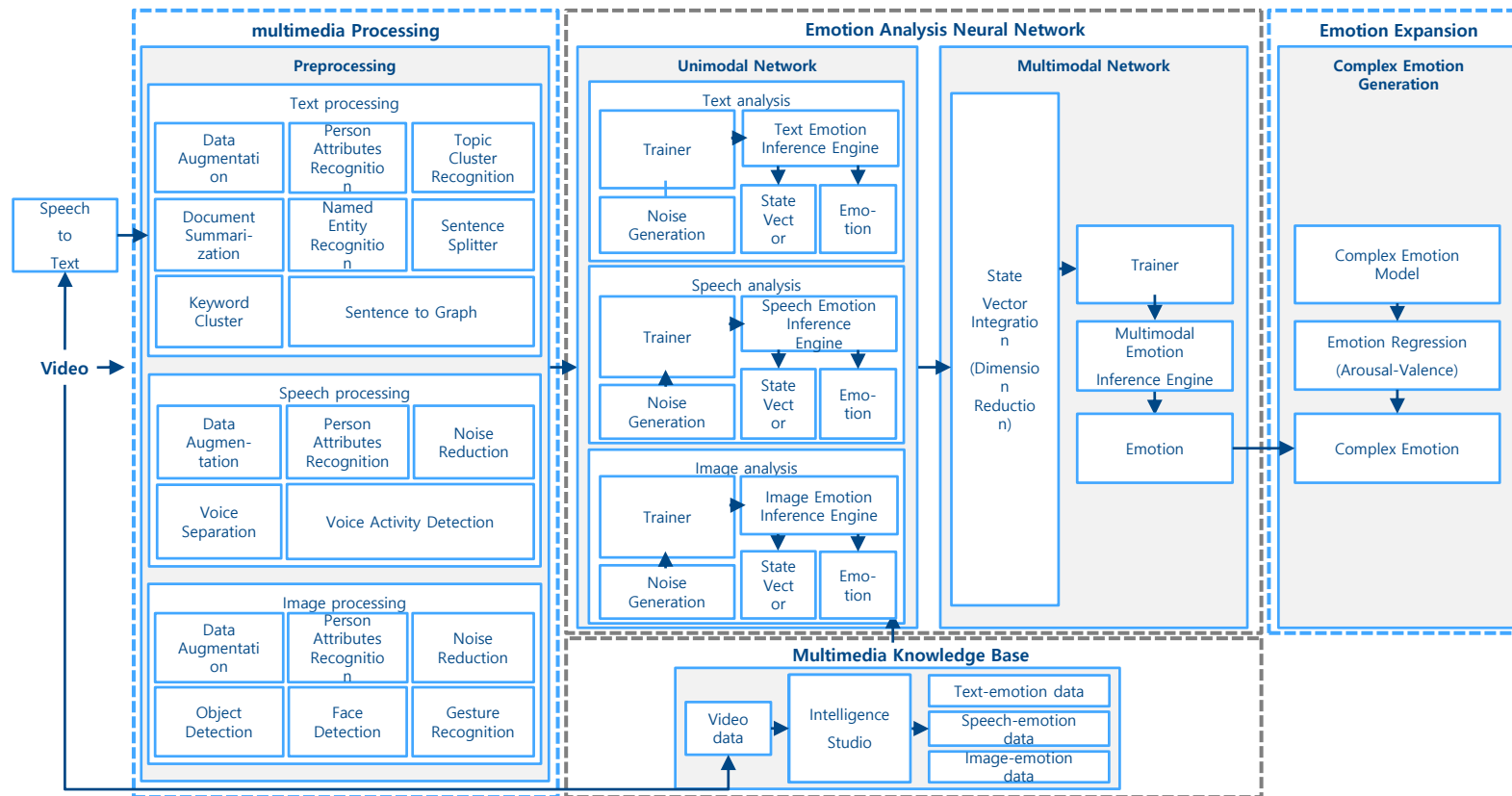
- E.FAST: User interface for face-to-face speech translation considering human factors
- E.OKID: On-screen Keyboards for ICT devices
- F.HFS-BC: Requirements and framework for blockchain-based human factor service models
- F.HF-SLM: Requirements and framework for ICT sleep management service models
- F.UI-SH: User interface requirements and framework for e-services based on speech/NLP technology
- F.EMO-NN "Emotion enabled multimodal user interface based on artificial neural network"
- BaaS.HC.req: Requirements of Blockchain as a Service (BaaS) for Healthcare Services with Human Factor
- BaaS.SM.PHR "Service Models of Blockchain as a Service (BaaS) for Personal Health Records (PHRs) with Human Factor"



## F.EMO-NN “Emotion enabled multimodal user interface based on artificial neural network” (아크릴과 ETRI 공동개발 표준)

- Emotion technology achieved big improvement in **HCI** (Human Computer Interaction)
- **Multimodal emotion technology** has become one of the most promising area as it uses same sources with human: language, sound, image.
- Various applications using **multimodality and emotion analysis**
- Study items for **emotion analysis** include:
  - Pre-processing methods for multimedia data
  - Integration method between unimodal and multimodal network
  - Expansion of **complex emotion** with additional knowledge

# Functional components and interfaces for multimodal emotion recognition service

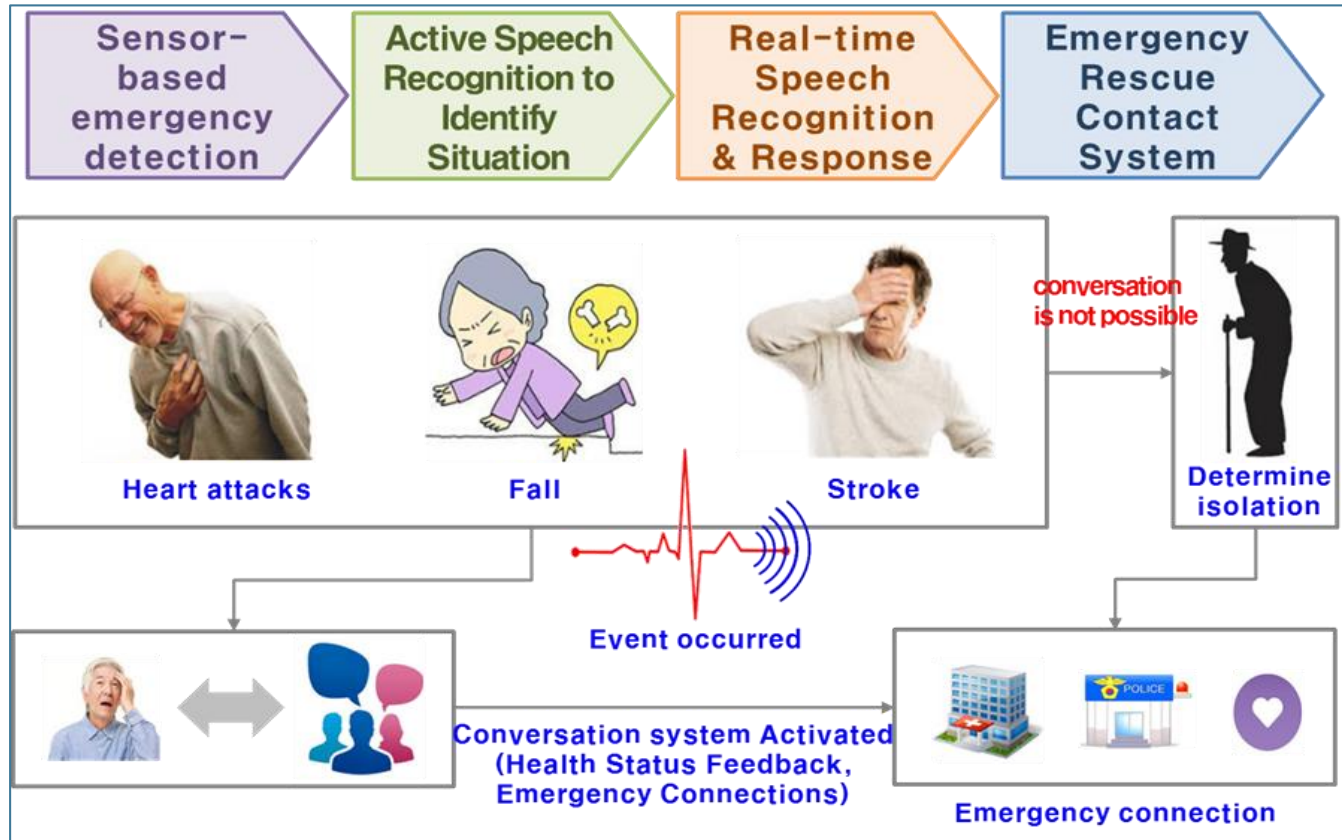






## F.UI-SH “User interface requirements and framework for e-services based on speech/NLP technology”

- The use of **speech recognition in traditional healthcare sector** has been used only as a means to enhance the convenience of medical personnel in producing medical records. However, recent speech recognition services and researches have improved the convenience of typing beyond simple user interface through natural dialogue with users
- In addition, it has become possible to manage customer's health, such as checking and improving their health status
- Various voice-based services are being developed from a nursing robot to care patients and to identify problems through conversation with patients, to the technology for early diagnosis of dementia from voice
- This contribution proposes a service model through the speech recognition interface in healthcare fields using several scenarios to apply speech recognition technology to healthcare services



Example of smart health services based on speech recognition interface



## F.HF-SLM “Requirements and framework for ICT sleep management service models”

- Korea is the country with the title of the lowest sleeping time in the OECD, and sleep will affect labour productivity, healthcare in many fields of human factors.
- While previous sleep management markets focused on furniture including bedding and medical sleep management, current trends include a new form of sleep management technology, including a convergence of existing sleep management methodology and ICT.
- These issues have led to the creation of a separate sleeping technology exhibition at CES last year.
- The most actively studied areas of sleep management services are:
  - Sleep monitoring (sleep monitoring device using various sensors)
  - Sleep improvement (sleep improvement device to enhance sleep quality)
  - Inducement of deep sleep (inducing deep sleep using lighting, white noise, etc.)



Example of sleep management services



## F.HF-BC “Requirements and framework for blockchain-based human factor service models”

- Recently, Blockchain technology has become a global issue through application of cryptocurrency (for example, Bitcoin)
- This contribution proposes a model for applying Blockchain to human factor services
- Human factor services are related issues for improvement of the quality of life including occupational health and safety, patient safety and quality improvement science, offering an integrated, evidenced and coherent approach to patient safety, quality improvement and clinical excellence



# BaaS.HC.req “Requirements of Blockchain as a Service (BaaS) for Healthcare Services with Human Factor”

- Blockchain technology is currently expanding its usefulness in various fields
- The healthcare sector is an area where sensitive information of individuals is distributed, and is recognized as an important area in applying and utilizing the blockchain technology.
- This recommendation defines the terms used in various healthcare services and defines the minimum requirements for blockchain technology based services
- Software as a Service (SaaS) refers to the ability of third parties to take software and use it as a service, such as remote cloud environments, without the ownership of software resources
- Blockchain as a Service (BaaS) is a blockchain used environment that takes and utilizes blockchain in a similar way to SaaS





# BaaS.SM.PHR “Service Models of Blockchain as a Service (BaaS) for Personal Health Records (PHRs) with Human Factor ”

- This recommendation describes how users can apply the blockchain service model to ensure that data is self-determined in order to ensure data reliability and interoperability in PHR services
- Personal health records are generated from various sources such as hospitals, health care services, and personal health devices
- It is difficult for individuals to make decisions on the use of the data and track/manage it.
- Software as a Service (SaaS) refers to the ability of third parties to take software and use it as a service, such as remote cloud environments, without the ownership of software resources
- Blockchain as a Service (BaaS) is a blockchain used environment that takes and utilizes blockchain in a similar way to SaaS



- Title: **Accessibility to multimedia systems and services**
- Motivation:
  - functional limitations, disabilities, or other natural causes.
  - many users will have sensory and motor limitations.
  - users can make benefit of the mainstream telecommunication services.
  - universal design of these services and systems, making them accessible to as many users as possible.



# ITU-T SG16 Q.26 접근성 그룹 (표준화 주제 1)

- Study items:
  - clauses on **accessibility issues in relevant Recommendations**, declaring how inclusive design is achieved
  - support for wide performance limits in production, perception and control of each medium in communication services to allow for maximum **usability** in accordance with the **principles of universal design**. Specifically, study profiling of the latest video coding standards to fulfil sign language and lip reading needs at very low bit rates and in error prone environments;
  - study potential **accessibility benefits** offered by **emerging technologies**, such as independent living, home automation, communication between smart things, cloud-based service and smart homes;
  - specification of interfaces on communication equipment to allow various forms of **user interface** equipment to be attached in order to enable session and device control and media handling by persons with varying capabilities and preferences;  
NOTE – Examples of what the interfaces should support include: talking menus, keyboards, pointing devices, listening and viewing devices, Braille and voice call control, text conversation input and output.



# ITU-T SG16 Q.26 접근성 그룹 (표준화 주제 2)

- Study items:
  - multimedia services with **multimodality** for capabilities and preferences of end users.
  - mechanisms for **user selectable media**, including its production, storage, transport, presentation and logical linking;
  - specification of accessible services using wireless telecommunication technologies
  - maintenance of the **total conversation concept**, and its inclusion in any new multimedia conversation protocol;
  - study access to **emergency services and early warning services** by persons with disabilities and specific needs with a wide range of communication channels, e.g. text, sign language, and lip reading supported speech, audio description, and braille;
  - study mechanisms for disability inclusive **disaster risk reduction**.



# ITU-T SG16 Q.26 접근성 그룹 (기존 표준 1)

- Recommendations: (real time text telephony)
  - ITU-T V.18 for real time text telephony;
  - ITU-T T.140 as the general presentation protocol for real time text conversation;
  - ITU-T T.134 for real time text conversation in the T.120 data conferencing environment;
  - Annex G to ITU-T H.323 for real time text conversation in H.323 packet multimedia environment;
  - Annex L to ITU-T H.324 for real time text conversation in low bit-rate multimedia applications;



# ITU-T SG16 Q.26 접근성 그룹 (기존 표준 2)

- Recommendations:
  - ITU-T F.790 – Telecommunications accessibility guidelines for older persons and persons with disabilities;
  - ITU-T F.791 – Accessibility Terms and Definitions
  - ITU-T H.702 – Accessibility Profiles for IPTV Systems
  - ITU-T Technical Paper FSTP-AM – Guidelines for Accessible Meetings
  - ITU-T Technical Paper FSTP-ACC-RemPart – Guidelines for supporting remote participation in meetings for all
  - ITU-T Technical Paper FSTP-TACL – Telecommunications Accessibility Checklist.



## Q.26 접근성 그룹의 표준화 현황 (진행 표준1)

- F.WAAD "Safety requirements for audio augmenting devices"
- FSTP.ACC-AM "Accessible meeting"
- FSTP.ACC-RemPart "Remote Participation"
- FSTP.Intl-Relay "International Relay Services"
- FSTP-TRS-KPI Technical Paper: "Key performance indicators for telecommunication relay services"
- H.702 "Accessibility Profiles for IPTV"
- H.ACC-GAD "Guidance on audio descriptions (New)" (twin text of ISO/IEC TS 20071-21:2015, Information technology - User interface component accessibility - Part 21)
- H.ACC-GAP "Guidance on the audio presentation of text in videos, including captions, subtitles and other on-screen text (New)" (twin text of ISO/IEC 20071-25:2017, Information Technology - User interface component accessibility Part 25)
- H.ACC-GVP "Guidance on the Visual presentation of audio information, including captions and subtitles" (twin text of ISO/IEC DIS 20071-23, Information technology - User Interface component accessibility Part 23)



## Q.26 접근성 그룹의 표준화 현황 (진행 표준2)

- H.ACC-MMSIGN "Abstract language for multimedia signing"
- H.ACC-RCAD "Requirements for captioning and audio description for accessibility"
- H.MD-DiDRR "Profile metadata for persons with specific needs as part of disability-inclusive **disaster risk reduction**"
- HSTP.ACC-AUD "Technical Paper on Methods for improving the intelligibility of audio (or speech)"
- HSTP.ACC-Interop "Interoperability of digital audiovisual media accessibility"
- HSTP.ACC-SL "Production guidelines for sign language service"
- HSTP.ACC-UC "Multimedia accessible system use cases"
- HSTP.AEHH "Audio enhancement for the hard-of-hearing"





# 기타 접근성 표준 활동 그룹

- ITU-T SG20 IoT **관련 접근성 표준화** (Q.2 Requirements, capabilities, and use cases across verticals )
  - **Y.Accessibility-IoT** "Accessibility requirements for the Internet of things applications and services"
  - provide accessibility requirements specific to IoT applications and services
- ITU-D Q.7 **개발도상국들에게 접근성 정보 제공, 4년마다 접근성 보고서 출간**
- ISO JTC1 SC35 User Interface **그룹의 접근성 표준화 활동**
  - WG6 "User interfaces accessibility"
    - Guidance on the visual presentation of audio information (including captions and subtitles)
    - User interface accessibility— Part 1: User accessibility needs
    - A code of practice for creating accessible ICT products and services
    - Guidance on text alternatives for images



# Thank You