

SW그랜드챌린지: 엑소브레인 SW 개발



2015. 8. 21.

박상규
SW기반기술연구본부
한국전자통신연구원

발표 순서

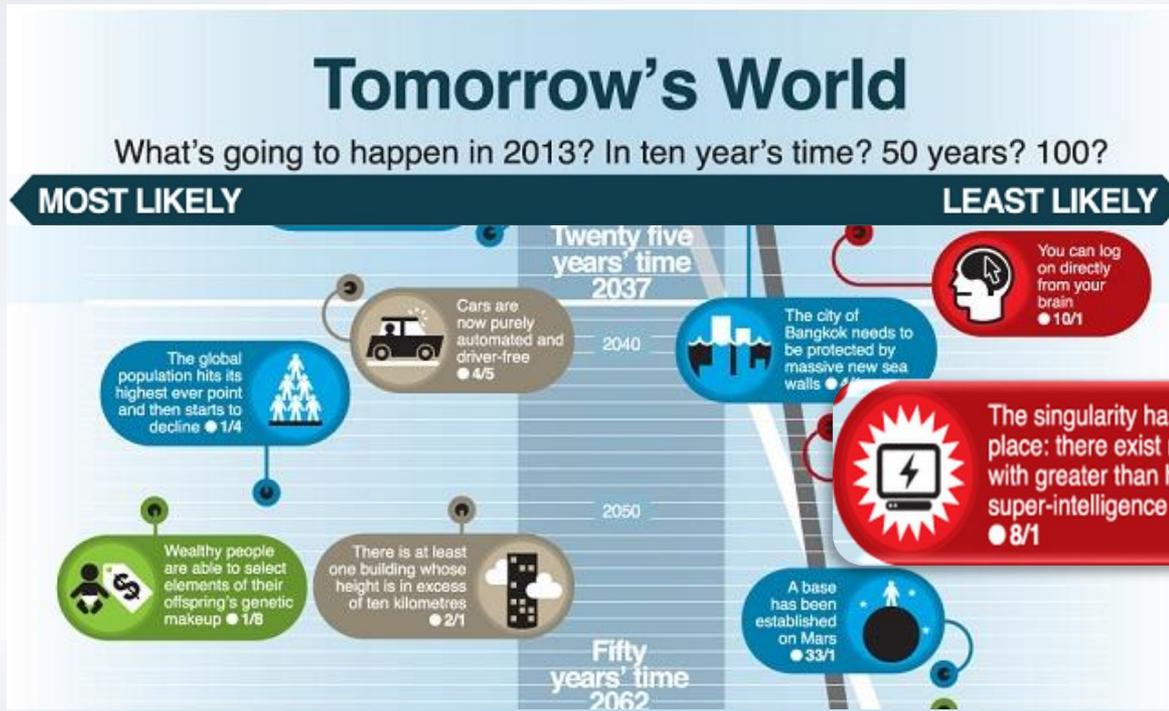
I 엑소브레인과 인공지능

II 엑소브레인 SW 추진성과

III 엑소브레인 SW 추진계획

IV 엑소브레인 SW 활용방안

- **AI 정의: Science and engineering of making intelligent machines**
 - * John McCarthy, 1955
- **특이점 전망: 전문가에 따라 의견이 다르나, 30년 ~ 1000년 이내로 전망**
 - * BBC(2013)와 유엔 미래보고서(2015) 등은 2045년 도래 예측
 - * 특이점: AI가 인간의 지능을 넘어서는 시점



Sources: ABC, BBC, CNN, Discovery News, the Guardian, IBM, Information Week, The Life Scientific, Long Bets, MIT Technology Review Report, NASA, Popular Mechanics, 'The Singularity is Near', Time magazine, Trendwatching, United Nations Environment Programme. Odds provided by Ladbrokes



〈출처: 유엔 미래보고서 2045(2015)〉

〈출처: Tomorrow's world: A guide to the next 150 years(BBC, 2013)〉

- Weak AI: Simulates human thoughts and actions
- Strong AI: Matches or exceeds human intelligence

데이터와 지식
(언어, 영상, 감각 등)



지능을 가지는
컴퓨터/머신



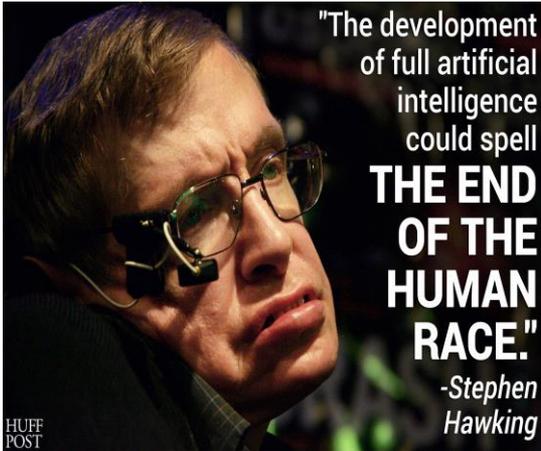
[언어지능] 인간의 말과 글을 사람처럼 이해하고 수행하기 위한 자연어 이해 및 지식화 처리 기술

[시각지능] 인간이 눈으로 들어오는 영상의 의미를 파악하듯 영상의 상황을 이해하고 예측하는 기술

[청각지능] 인간이 자연스럽게 구사하는 발성을 문자로 변환하는 비정형 자연어 대화체 음성인식 기술

[뇌 인지컴퓨팅] 인간 두뇌의 생물학적 특징을 모델링 하여, 인간의 사고 체계를 접근하는 기술

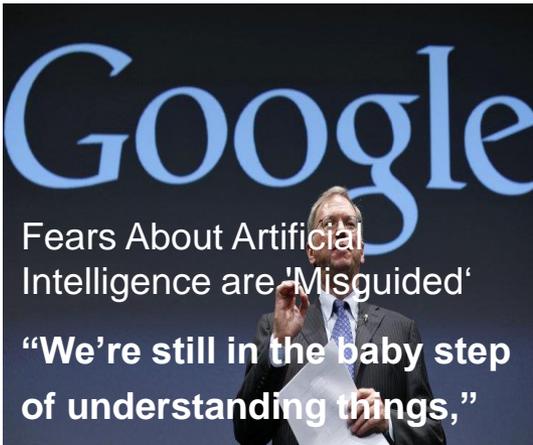
개발 동향



• IBM, 구글, MS 등 글로벌 기업은 경쟁적 개발 시작

기업	로고	내용
IBM		'왓슨' 슈퍼컴퓨터 중심으로 생태계 조성
구글		무인자동차, 얼굴과 음성 인식 등 인공지능 적용 '딥마인드' 등 각종 업체 인수합병 활발
MS		음성 인식 갖춘 개인비서 '코타나' 올해 출시
페이스북		인공지능 연구그룹 구성, 얼굴 인식 프로그램 발표
애플		'시리' 업그레이드 버전으로 지능형 개인비서 개발

기술 수준

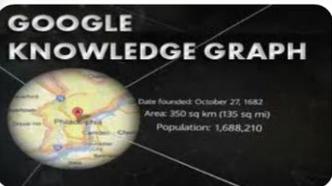


- 딥러닝 기술 적용으로 음성과 이미지 인식 성능 향상
 - 얼굴인식률: 구글 '페이스넷(FaceNet)' 99.96%
- 사물의 이해 측면에서 인공지능은 아직 초보적 수준
 - 출처: Google's Eric Schmidt (2015)
- Weak AI 위주 시장 예측(Tractica, 2015)
 - Weak AI is gaining ground faster than Strong AI.



Watson Deep QA 기술개발로 퀴즈쇼에서 인간을 능가(2011)

- 퀴즈분야 특화 단답형 응답 한계. 헬스케어 및 파이낸스 분야 확장 중



5억개 Knowledge Graph를 구축하여 짧은 질의에 대한 요약형 정보 제공

- 구조정보 기반 인물, 지역, 사물 정보 위주 제한된 정보 제공



후지쯔와 NII는 2021년 동경대 입시 합격을 목표로 프로젝트 진행

- 토다이 로봇 프로젝트: 슈퍼컴퓨터 활용 10년(2011~2021)간 진행



미래부에서 2013년 SW 그랜드 챌린지 과제로 엑소브레인 SW 과제 시작

- 단기적 R&D 한계를 극복한 지속적 버전업 R&D 수행 시 시너지 기대

* QA(Question Answering): 자연어 질문에 대한 정답 제공 기술
* NII(National Institute of Informatics): 일본 국립정보학연구소

발표 순서

I 엑소브레인과 인공지능

II 엑소브레인 SW 추진성과

III 엑소브레인 SW 추진계획

IV 엑소브레인 SW 활용방안

사람과의 의사소통을 뛰어넘어 지식 소통이 가능한 인공지능 SW 개발

* 지식 산업환경에서 전문가 수준의 질의응답(Question Answering)을 통한 지식 서비스 제공

Question Answering

- 교육/법률/특허 등 전문지식 상담

무배당 신바람 건강보험에 가입했습니다. 신장결석으로 충격파쇄석수술을 했는데, 수술시 보험금 지급이 가능한가요?



자연어 심층이해

Decision Support

- 지식 기반의 판단 지원

애플 아이워치 출시가 국내 기업에 미칠 수 있는 영향은?



자연어 지식 생산

Smart Advisor

- 전문직종 의사결정 지원

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



엑소브레인

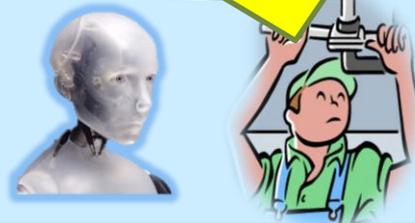


자연어 질의응답

Thinking Machine

- 지능형 단말의 인공두뇌로 활용

착륙예정점에 대한 경사각 수직면 유도 결합의 정비 항목은?



도전적 원천기술 확보로 시작, 글로벌 기술 상용화로 발전

2013.05

2017.02

2020.02

2023.02

목표
및
결과물

기술검증
및
사업화

1단계 (핵심 기술 개발)

IBM 왓슨과의 기술격차
7년을 4년간 개발하여 단축
(사실지식 QA 시스템)

검증방법: 인간과
엑소브레인과의
지식 대결 콘테스트



지식기반 QA 프로토타입
(예: VOC 분야)

지식기반 QA



분석 리포트 제공

2단계 (응용 기술 개발)

전문지식 서비스
세계적 성능 달성
(Smart Advisor)

전문지식 QA 시스템 상용화
(예: 의료, 법률, 금융, 상담분야)

지능형 콜센터 QA



헬프데스크 상담 QA

전문의 보조 QA



진단, 건강 보조

전문분야 컨설팅 QA



특허/경영/법률 자문

3단계 (글로벌 기술 개발)

다국어 전문지식 서비스
세계 최고 성능 달성
(Thinking Machine)

의사결정 지원 시스템 상용화

다국어 지식 QA



다국어 지식서비스 제공

지능형 로봇 QA



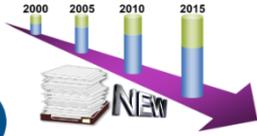
정비보조 등 지능적 업무 수행

웨어러블 QA



헬스케어, 피트니스 지식 서비스

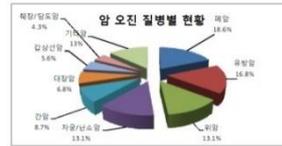
의료 환경 진단



새로운 의료 정보:
매 5년간 2배씩 증가



지식획득시간 절대부족
현실 : 5시간/1달



암 진단 오진율:
44% 증가 (2012년)



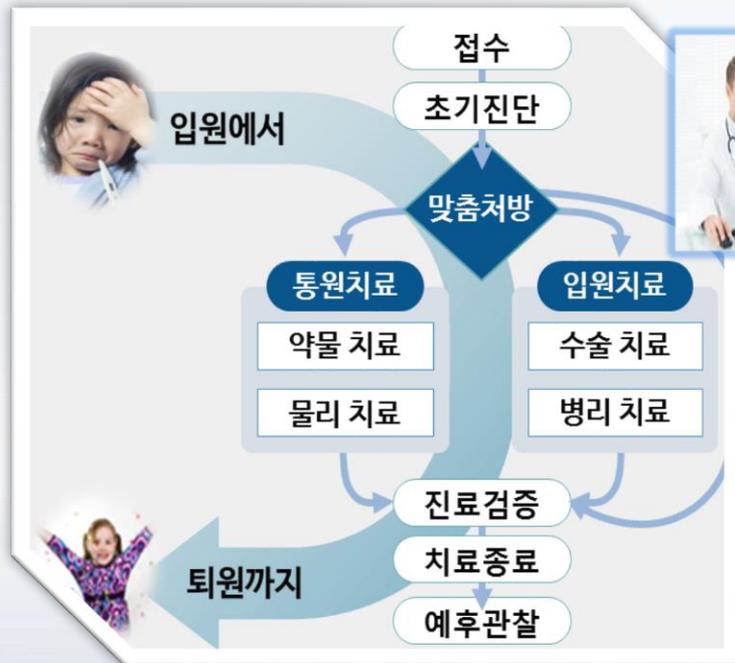
5대질환 사망자:
44% 증가



암 관련 의료비용:
42% 증가

최고, 최신의 대안 제시로 오진 예방

- 빅데이터 분석 기반 대안 진단 및 처방 추천
- EMR, 의학저널, 임상실험, 제약정보 등 분석
- 의사의 대안 검증을 통한 오진 및 의료비용 감소



질문

응답



Dr. 엑소브레인

환자 병력 및 빅데이터 분석

대안 진단 제시

대안 처방 제시

환자 맞춤형 솔루션 제시

의료 지식베이스

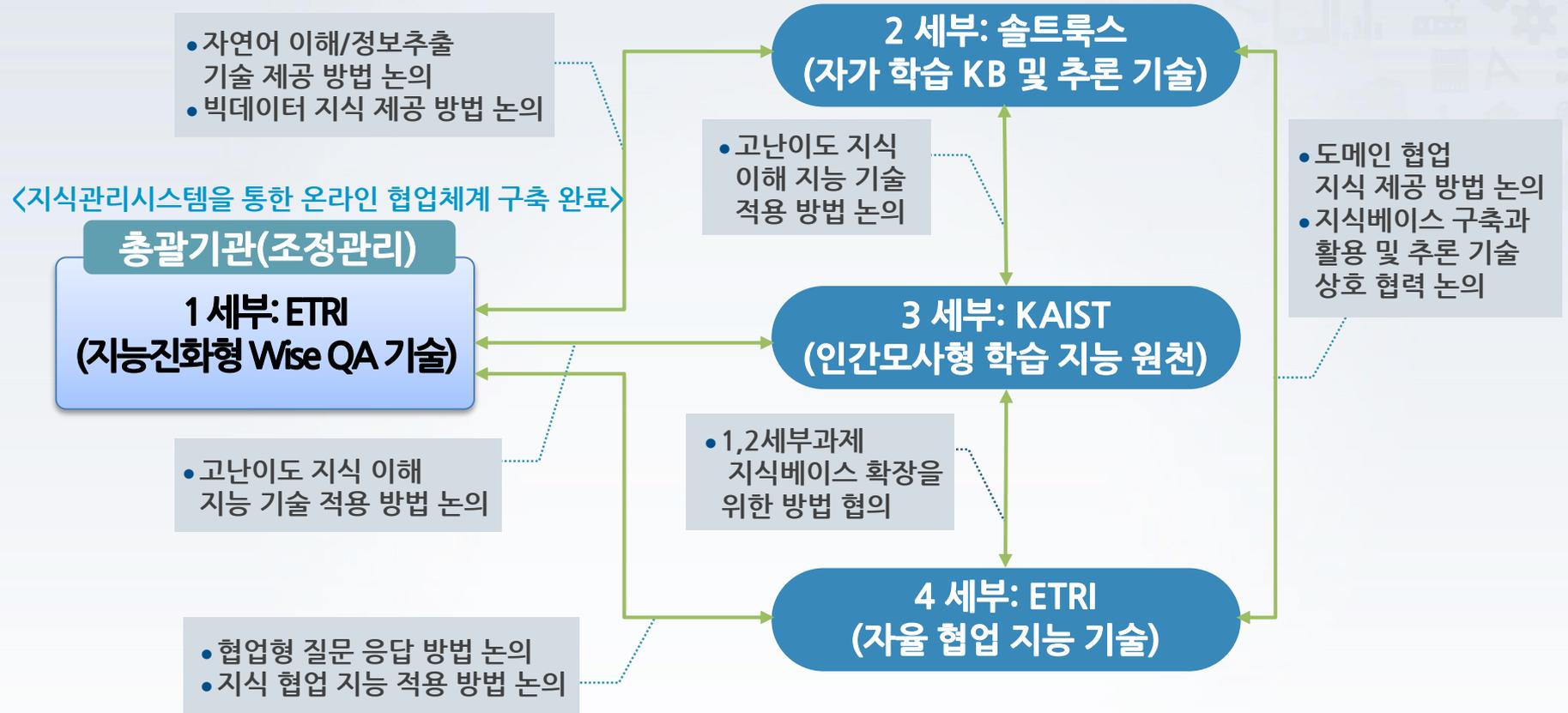


협업 에이전트
(A 중합병원)



협업 에이전트
(해외 암 센터)

세부과제 간 긴밀한 연계 조정, 협력 체계 마련



〈세부과제 협력 추진〉

- ◆ 정기적 과제 워크숍 및 운영위원회 개최
- ◆ 매월 과제책임자, 실무책임자 회의 개최
- ◆ 공통 태스크 해결을 위한 세부과제 간 TFT 추진

국내 최고 인공지능 산/학/연 컨소시엄을 통한 생태계 확산 추진

출연연

- ETRI 주도로 자연어 이해 및 질의응답 원천 기술 개발
- 세부과제 총괄 역할 수행 및 인공지능 산/학/연 생태계 조성

산업체

- 솔트룩스 주도로 나눔기술 등 11개 기업 참여
- 지식베이스 구축 기술 개발 및 상용화 기술 개발

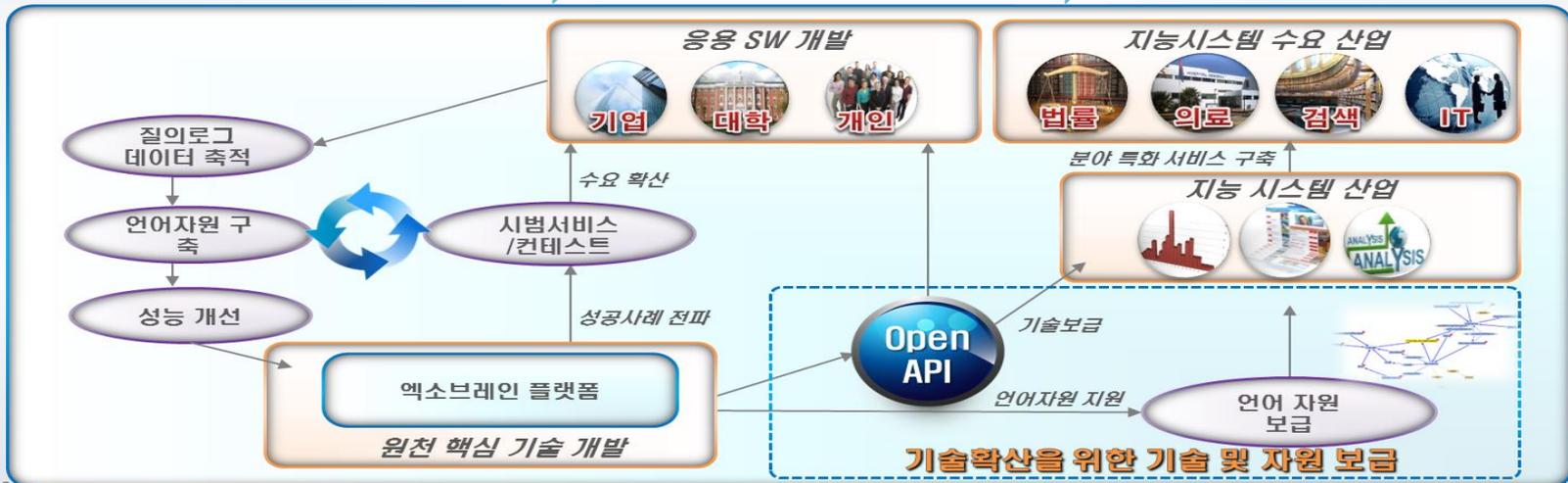
대학

- KAIST 주도로 서울대, 포스텍 등 15개 대학 참여
- 인간 모사형 자가학습 및 지식 추론 기술 연구

인공지능 생태계 조성 및 확산

산업 생태계 조성

글로벌 생태계 조성



인공지능 원천기술 개발

1

세계적 수준 자연어 이해 시스템 개발

자연어 구문분석 성능: 엑소브레인 91.2% vs. 왓슨 88.7%

2

아시아 최대 지식베이스 구축 추진 중

한국어 위키백과 7배 규모 (4,480만 트리플) 지식베이스 구축

3

복합 추론형 QA 기술 검증 완료

귀납적/연역적 복합 추론을 통한 QA 방법론 정립 및 프로토타입을 통한 검증 완료

품질관리, 생태계, 표준화 추진

4

글로벌 기업 (IBM, NASA) 수준 SW 품질관리 인증 획득

ISO SPICE Level 3 품질관리 인증 획득('14.12.5.)

5

인공지능 산/학/연 생태계 확장 추진

SW 배포(7개 기술, 14개 기관), 국제워크샵 5회, 워크샵 및 교육 22회

6

인공지능 기술의 IPR 확보

특허 출원 106건, 논문 게재 177건(SCI/E 30건), 표준화 기고서 18건
ITU-T 국제표준 채택 예정: Intelligent QA Service Framework('15.12)

발표 순서

I 인공지능 개요

II 엑소브레인 SW 추진성과

III 엑소브레인 SW 추진계획

IV 엑소브레인 SW 활용방안

1

장학퀴즈 우승자 수준의 엑소브레인 조기 개발

2

WiseQA 기술의 국제 표준안 채택 추진 (국제특허와 연계 IPR 확보)

3

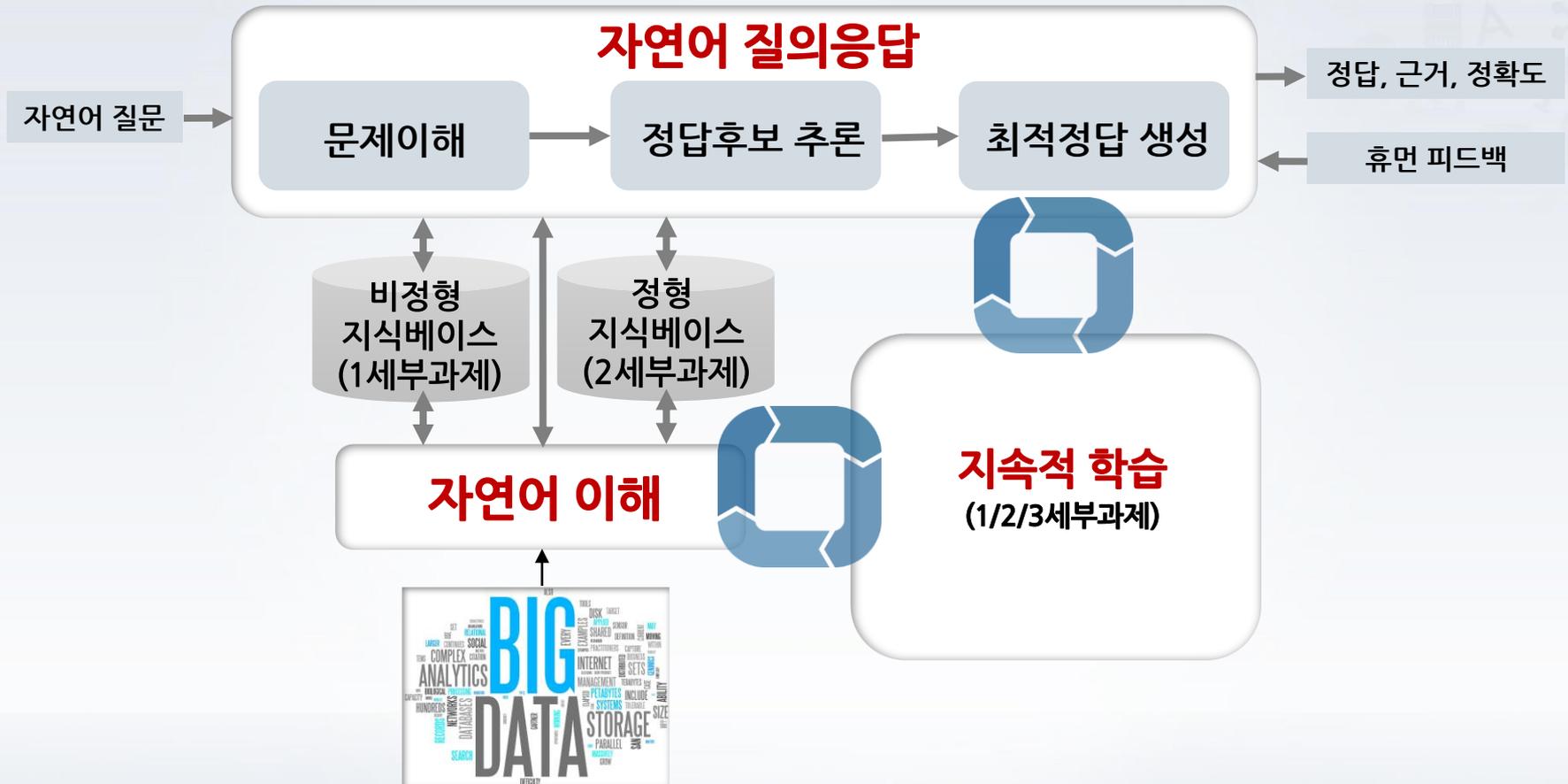
인공지능 생태계 활성화 및 사업화 모델 발굴

인간과 엑소브레인과의
지식 대결 콘테스트
(2016년 예정)

1

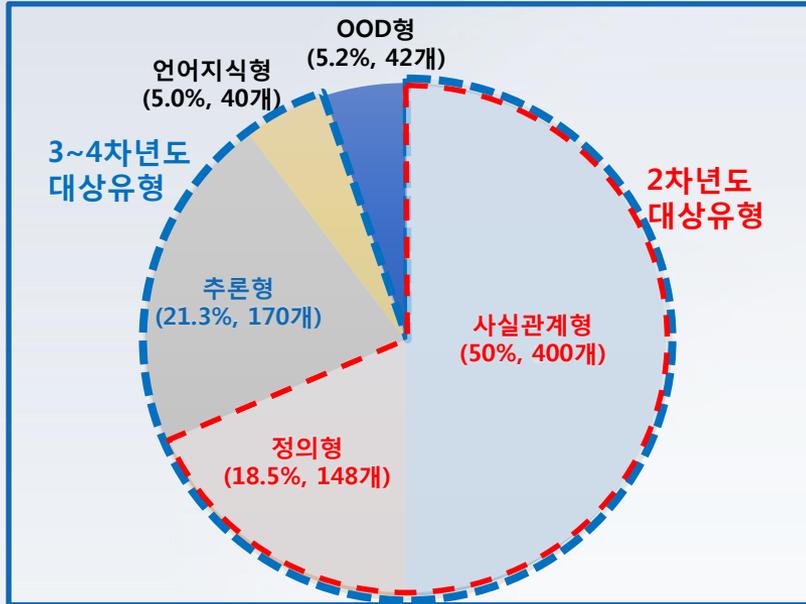
장학퀴즈 우승자 수준의 엑소브레인 기술의 조기 개발

정확률@응답률: 90%@50%, 응답속도: 10초 이내



WiseQA 문제 난이도 분석 (IBM Watson 대비)

질문 유형 분포(장학퀴즈 800문제)



IBM Watson과의 비교

질문 유형	Watson	WiseQA	차이
정의형	14.2%	18.5%	+4.3%
사실관계형	71.6%	50.0%	-21.6%
언어지식형	7.4%	5.0%	-2.4%
추론형	4.5%	21.3%	+16.8%
Out-of-Domain형	2.3%	5.2%	+2.9%

추론형 문제비율 높음
Watson 대비

- **정의형 (Definition)**
 - 세포 안에는 여러 가지 소기관이 있다. 식물에서 광합성을 담당하는 것은 엽록체인데, 그렇다면 우리 몸의 세포 안에서 호흡을 통해 필요한 에너지를 만드는 이 기관은 무엇일까? (미토콘드리아)
- **사실관계형 (Category-Relation, Number, Multiple-Choice, Date)**
 - 19세기 말 제국주의에 의한 식민지 쟁탈전이 치열해지는 가운데 아프리카에서 영국의 종단정책과 프랑스의 횡단정책이 충돌한 사건은 무엇일까? (파쇼다 사건)
- **언어지식형 (Abbreviation, Etymology, Translation, 사자성어, 속담)**
 - '시간이 상당히 지나는 동안'을 뜻하는 '한참'은 이것과 이것 사이의 거리를 뜻하는 말에서 유래했다. 이것은 무엇일까? (역)
- **추론형 (Fill-in-the-Blank, Common Bond, List, Calculation)**
 - 정지상의 한시<송인>, 고려가요<서경별곡>에서 공통된 이별의 장소로 등장하는 강은 무엇인가? (대동강)
- **Out-of-Domain형 (그림, 음악 등 시청각 문제)**
 - 이 그림은 크기와 모양이 같은 상자의 앞과 위를 본 것입니다. 이때 네모상자의 최대 개수는 몇 개일까요?

지식대결 도전과제 우승자 분석 (예: 장학퀴즈)

- 장학퀴즈 우승자 분석: 86.7%@45.2%, 14.2초 이내
 - 10회 (901회 ~ 910회) 분석 결과

라운드	정확률	응답률	최초 응답시간 및 정답비율	정답 응답시간	비고
1	우승자: 90.7% 참가자: 74.8%	우승자: 46.7% 참가자: 31.8%	6.8초 78%	우승자: 9.7초 참가자: 9.5초	<ul style="list-style-type: none"> 주관식 15문제 질문 길이: 66자
2	우승자: 77.7% 참가자: 58.7%	우승자: 42.5% 참가자: 36.4%	13.4초 62.6%	우승자: 18.6초 참가자: 17.1초	<ul style="list-style-type: none"> 객관식 30문제 질문 길이: 68자
3	우승자: 91.6% 참가자: 65.3%	우승자: 46.3% 참가자: 32.6%	12.4초 58.3%	우승자: 14.4초 참가자: 16.8초	<ul style="list-style-type: none"> 종합형 5문제, 최종 1문제 질문 길이: 97자

- 최초에 오답을 응답한 비율이 예상외로 많음: 참가자 33.7%, 우승자 15%
- 우승자의 평균 응답시간이 참가자보다 빠르지 않음: 3라운드 제외

- 장학퀴즈 우승을 위한 최소 성능: 90%@50%, 10초 이내 (80%@80% 성능과 유사)
 - 높은 정확성, 중간 정도 응답률
 - 라운드 별 퀴즈 전략 필요

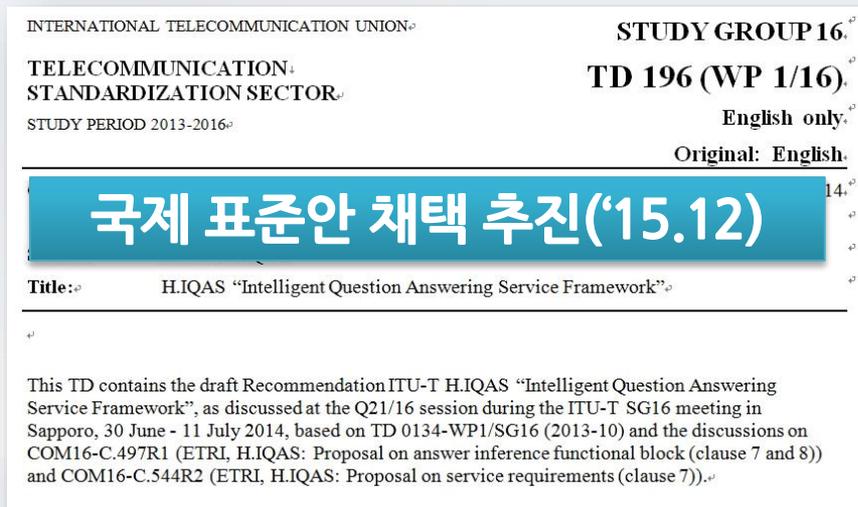
2 WiseQA 기술의 국제 표준안 채택 추진 (국제특허와 연계 IPR 확보)

● 국제 표준화

- ITU-T “Intelligent QA Service Framework” 국제 표준안 채택 추진
- 국제특허를 포함한 기고를 통해 표준 완료 이후 IPR을 통한 수익 창출 추진

● 국내 표준화

- TTA “형태소 품사 태그세트” 국내 표준안 개정 완료(2015.6.23.)
- 개체명 태그세트, 의존구문분석 태그세트, 의미역 태그세트 국내표준화 추진 예정



〈ITU-T 지능형 QA 서비스 프레임워크 기고서〉



국내 표준안 개정 추진('15.8)

〈TTA 형태소 품사 태그세트 기고서〉

엑소브레인 인공지능 연구 생태계 활성화

● 추진 목적

- 엑소브레인 인공지능 SW 국내 보급을 통해 국내 산업 경쟁력 강화, 상용화/사업화 추진
- 자연어 이해, 심층 기계학습, 자연어 질의응답 등의 차세대 유망 인공지능 분야 연구인력 양성

● 추진 전략



01

- 엑소브레인 SW 배포 (연구용도 한정)
- 엑소브레인 말뭉치, 평가셋, 지식베이스 배포

02

- 엑소브레인, 인공지능 인력양성 교육/세미나
- 공개 워크샵, 기술 세미나 추진

03

- 엑소브레인 연구성과에 대한 언론 홍보 (국제표준화, 도전과제 추진 등)

[언론 보도] 'IBM 왓슨 대항마를 키워라' ... 인공지능 전쟁 확산

- 미래부 2차관, 인공지능 SW 연구현장 방문 및 전문가 간담회 개최('15.6.10)
 - 정부, '왓슨 추격전' 기술확보 총력
 - 엑소브레인 프로젝트, IBM 왓슨과의 기술격차 7년을 1단계에 4년으로 단축 목표
 - 자연어 구문 분석 성능은 91.2%로, 왓슨(88.7%)을 근소한 차이로 넘어서
 - 한국어 위키백과의 7배 규모에 달하는 지식베이스를 구축 중, 복합 추론기능 검증
 - 오는 2018년 부터는 왓슨을 넘어서겠다 (목표)
 - 전문지식(예: 의료, 법률 등)에 대한 질의응답 서비스 제공 예정



최재유 미래부 제2차관, 인공지능 SW 연구현장 방문 & 간담회 (장소: 솔트룩스)

발표 순서

I 엑소브레인과 인공지능

II 엑소브레인 SW 추진성과

III 엑소브레인 SW 추진계획

IV 엑소브레인 SW 활용방안

인공지능, Deep Learning, 스마트머신 → **자연어 질의응답 기술**

* 인공지능 분야의 제일 어려운 난제이나, 성공시 매우 수익성이 높은 미개척 분야

IBM 왓슨

왓슨QA기술의 헬스케어, 금융분야 사업화 추진 중

토다이 로봇

“2021년에 도쿄대 꼭 합격할 거야“

셸

고객 안내 데스크에 인공지능 “아멜리아“ 얹힌다

유진 구스트만

13세 수준으로 대화, Turing Test 최초 통과

스탠포드 대학교

인공지능 백년대계 “AI 100” 프로젝트 착수

유엔미래보고서2045 인공지능: 더 이상 예측할 수 없는 미래가 온다

- ‘UN 미래보고서 2045’

- 인공지능의 발달이 고도화되는 미래에는 금융컨설턴트, 은행원, 콜센터 직원 등 대부분의 금융업 관련 직종이 인공지능에 의해 대체될 것
- 과거 IT기술은 비용 절감의 도구로 많이 활용되어 노동력을 대체하는 모습을 보였으나, 현재의 인공지능 기술 수준은 **노동력을 더 효율적으로 보완해줄 수 있는 방향으로 활용 가능**

- 스탠포드 대학 AI100(인공지능 백년대계) 프로젝트

‘인공지능 100년 연구 프로젝트’의 주요 연구 분야 내용

 기술 트렌드 인공지능이 인간의 지능을 이길 수 있을까	 경제 주식시장 애널리스트는 인공지능으로 대체될까. 인공지능은 미래 직업을 어떻게 바꿀까.
 사생활 침해 인간을 감시·통제하는 ‘빅브러더’가 현실화될까. 나온다면 어떤 대책이 있을까.	 전쟁·안보 공격 목표를 스스로 결정하는 첨단무기가 나오면 전쟁은 어떻게 달라질까
 법률 인공지능의 판단 착오로 피해 발생 시 누가 책임지나. 인공지능을 자연인·법인(法人)에 이은 제3의 인격체로 규정할 수 있을까.	 인공지능에 대한 통제력 강화 인간의 기대와는 별개로 움직이는 ‘초(超)인공지능’이 나올까. 나온다면 어떻게 해야 할까.
 윤리 어떤 분야에서 인공지능 활용을 금지해야 할까	

자료: 미 스탠퍼드대

자연어 QA 원천기술 조기 개발을 통한 인공지능 新산업과 고용 창출

2단계: 전통산업 ICT化

전문 지식 제공

<p>빅데이터 QA</p> <p>빅데이터 분석 리포트 제공</p>	<p>콜센터 QA</p> <p>헬프데스크 상담 QA</p>	<p>전문가추천 QA</p> <p>전문분야 전문가 추천</p>
--------------------------------------	----------------------------------	------------------------------------

Question Answering

3단계: 지능 新산업 창출

스마트카, 디지털사이니지, 휴머노이드 로봇등에 탑재

<p>쇼핑센터 안내 아바타</p> <p>쇼핑센터 가이드</p>	<p>스마트카 QA</p> <p>위치, 장소 정보제공</p>	<p>지능형 로봇</p> <p>정비보조 등 지능적 업무 수행</p>
------------------------------------	-----------------------------------	---------------------------------------

Embodied Agents

엑소브레인
인공두뇌

대화형 스마트 컨설턴트

<p>전문 의 보조</p> <p>진단, 건강 보조</p>	<p>전문 지식 제공</p> <p>특허/논문 전문 지식 제공</p>	<p>전문 분야 컨설팅</p> <p>법률 자문</p>
---------------------------------	---------------------------------------	-------------------------------

Cooperative Partner

웨어러블 디바이스 환경

<p>피트니스 QA</p> <p>휴대형 디바이스</p>	<p>인포테인먼트 QA</p> <p>신체부착형 디바이스</p>	<p>헬스케어 QA</p> <p>생체이식형 디바이스</p>
--------------------------------	------------------------------------	----------------------------------

IoT Intelligence

국가.산업에 활용 가능한 자연어 QA 인공지능 개발

스마트 머신
(스마트 자문가: 인간의
지적노동 보조)

Top 10 Strategic Technology Trends for 2014
(Gartner, 2013)

지적 노동의 자동화
(5~7조달러 경제적가치)

Disruptive technologies
(McKinsey, 2013)

2011

2020

2024

2030

2045

(Techcast, 2013)

Tomorrow's world: A guide to the next 150 years
(BBC, 2013)

인공지능 재조명

**기계가 인간의
지적노동 30% 대체**

**인간지능을 뛰어넘는
슈퍼지능머신**
(확률 10%~50%)



감사합니다.