

빅데이터 지식처리 인공지능 기술동향

Technology Trends of AI for Big Data Knowledge Processing

이형직 (H.J. Lee)	휴먼컴퓨팅연구실 선임연구원
류법모 (P.M. Ryu)	지식마이닝연구실 선임연구원
임수종 (S.J. Lim)	지식마이닝연구실 책임연구원
장명길 (M.K. Jang)	지식마이닝연구실 책임연구원
김현기 (H.K. Kim)	지식마이닝연구실 실장

2014
Electronics and
Telecommunications
Trends

소프트웨어 기술동향 특집

- I. 머리말
- II. 기술개념
- III. 기술동향
- IV. 산업동향
- V. 지식재산권 동향
- VI. 발전방향
- VII. 맺음말

최근의 플랫폼 기술동향은 웹 기반 혹은 단순 의사소통이 가능한 모바일 플랫폼에서 빅데이터와 인공지능기술이 접목되면서 심층 질의응답이 가능한 차세대 지능형 지식처리 플랫폼으로의 진화가 진행 중이다. 선진국에서는 국가 차원 혹은 글로벌 기업의 주도하에 대형장기 프로젝트가 진행 중이다. 국가 주도의 프로젝트로는 미국의 PAL, 유럽의 Human Brain, 일본의 Todai 프로젝트가 대표적인 예이며, 글로벌 기업의 경우는 IBM의 Watson, Google의 Knowledge Graph, Apple의 Siri가 대표적인 예이다. 본고에서는 차세대 지능형 플랫폼의 핵심기술인 인간과 기계의 지식소통을 위한 빅데이터 기반의 지식처리 인공지능 소프트웨어 기술의 개념과 국내외 기술 및 산업, 지식재산권 동향 등을 살펴보고 산업계 활용 방안 및 발전방향에 대해 논하고자 한다.

I. 머리말

과거 HP, Sun, IBM 등의 서버업체들이 주도한 1세대 패러다임의 플랫폼은 Microsoft의 등장으로 데스크톱 기반의 PC가 주를 이루는 2세대 패러다임의 플랫폼으로 변화되었다. 최근 Google, Apple, Facebook 등의 웹 및 모바일 기반 플랫폼의 등장으로 다시 한번 플랫폼의 패러다임이 변경되었으며, 제 4세대 플랫폼의 패러다임은 방대한 데이터로부터 고수준의 지식창출이 가능한 인공지능 기반의 지능형 플랫폼으로 발전될 것으로 전망되고 있다. 이미 글로벌 기업들은 각자 지능형 플랫폼에 대한 연구를 수행 중이며 일부 기업들에서는 사용자를 이해하고 의도에 맞는 지식을 제공하는 서비스를 제공하거나 상용화하기도 하였다. IBM의 Watson, Apple의 Siri(Speech Interpretation and Recognition Interface), Google의 Knowledge Graph가 대표적인 사례이다. 이처럼 차세대 지능형 소프트웨어에 대한 소리 없는 전쟁은 이미 시작되었으며 지식 및 지능처리 기술을 선점하는 자가 이 전쟁에서 승리자가 될 것이며, 향후 IT 시장을 주도하게 될 것이다[1][2]. (그림 1)은 플랫폼 패러다임의 변화와 전망에 대해 나타내고 있다.

이에 본고에서는 지식 및 지능처리 소프트웨어의 핵심기술의 개념에 대해서 살펴보고 각 국가별 및 기업별 기술동향과 이 기술의 활용방안 및 발전방향에 대해서

살펴보겠다.

가트너와 IDC에서는 2013년 IT 기술의 메가트랜드 중의 하나로 빅데이터 기반의 인공지능기술을 선정하였으며, MIT에서는 추론과 학습 및 의사소통이 가능한 인공지능기술(Deep Learning)을 2013년 10대 돌파형 기술로 선정하였다. 미국에서는 IBM이 슈퍼컴퓨터 Watson을 이용하여 인공지능과 빅데이터 분석 소프트웨어를 접목시켜 질의응답에 최적화하여 ‘Jeopardy! 퀴즈쇼’에서 인간 챔피언을 물리치고 우승하였으며, 의료 및 금융 쪽으로 지식영역을 확장하고 있다. 일본에서도 2021년 동경대학교 입학이 가능한 수준의 인공지능을 갖춘 Todai 로봇 프로젝트를 수행하고 있다[3]. 유럽에서도 Human Brain 프로젝트를 통해 인간의 인지방식과 유사한 자율학습 모델에 대해 연구를 진행 중이다. 국내에서도 국가혁신형 과제인 엑소브레인 SW 기술개발 과제를 통해 선진국들과의 지능형 플랫폼 기술개발에 대한 경쟁을 시작하였다.

빅데이터 지식처리 인공지능 SW의 핵심기술은 관련 기술의 개념 부분에서 자세히 소개하겠지만 간략하게 다양한 기계학습 기술을 통해 자연어를 이해하고 표현하여 지식베이스를 구축하고 새로운 지식을 추론 및 생성해낼 수 있으며, 학습과 추론을 통해 진화가 가능한 인공지능기술로 요약할 수 있다.

II. 기술개념

빅데이터 지식처리 인공지능의 개념은 빅데이터로부터 스스로 학습하고 지식을 축적하며 사용자와 의사소통을 하고 필요에 따라 자율협업을 통해 지식의 공유 및 진화가 가능한 차세대 SW 기술을 의미한다. 빅데이터 지식처리 인공지능 SW는 첫째로 일반지식, 전문지식 등에 대한 분석/추론 및 심층학습을 통하여 전문가 수준의 문제 해결 및 의사결정을 지원할 수 있어야 한다.



(그림 1) 플랫폼 패러다임의 변화 및 전망



(그림 2) 빅데이터 지식처리 인공지능 SW의 개념도

둘째로 인간 수준의 지식·지능 체계를 가지고 의사소통 및 자가학습과 도메인 확장이 가능한 지식베이스를 스스로 구축하는 지능 진화형의 지식 생산능력이 있어야 한다. 셋째로 다양한 기기에 탑재되어 분산된 각종 지식 베이스 및 기기 간의 자율협업을 기반으로 발현되는 협업지능을 통해 새로운 문제를 스스로 사고하고 해결하는 기능을 가져야 한다. (그림 2)에 빅데이터 지식처리 인공지능기술의 개념도를 나타내었다.

III. 기술동향

국외는 주로 정부 주도 혹은 글로벌 기업의 주도로 지식/지능처리 소프트웨어 관련 연구가 진행되고 있으며, 국내는 국외 기업들과의 기술격차를 줄이기 위해 정부 주도로 장기간의 국책사업이 시작단계에 있다. <표 1>은 인공지능 및 심층 질의응답 기술의 동향을 나타낸다. 미국의 경우 정부의 주도하에 대규모의 장기 프로젝트

를 진행 중이다. Siri 서비스의 모태가 된 인공지능 분야의 PAL(Personal Assistant that Learns) 프로젝트와 심층 질의응답 시스템을 연구하는 AQUAINT(Advanced Question Answering for Intelligence) 프로젝트를 수행 중이며, 2012년 3월 연간 2억불 규모의 Big Data R&D Initiative 프로그램을 진행 중이다.

일본의 경우 2011년부터 Fujitsu 연구소와 NII(National Institute for Informatics)가 공동으로 2021년 동경대학교 입시 합격이 가능한 수준의 인공지능 시스템 개발을 위해 Todai 프로젝트를 진행 중이다.

IBM, Google, Apple 등의 글로벌 기업들은 심층 질의응답 기술이 차별화된 지능화 서비스를 주도할 핵심기술로 전망하고 있다. IBM은 DeepQA 프로젝트를 통해 Watson 질의응답 시스템을 2006년부터 개발하였으며, 2011년 2월 'Jeopardy!' 퀴즈 쇼'에서 역대 챔피언 2명을 물리치고 우승을 차지하였으며, 이후 의료, 금융 등의 전문분야로 확장하고 있다[4][5]. Apple은 음성인식 기

〈표 1〉 인공지능 및 심층 질의응답 프로젝트 현황

프로젝트명	수행기관 및 기간	주요 내용
Watson DeepQA	IBM 2006 ~ 2011	<ul style="list-style-type: none"> Factoid 형태의 복잡한 신뢰성 높은 빠른 응답 제시 http://researchweb.watson.ibm.com/deepqa/
Knowledge Graph	Google	<ul style="list-style-type: none"> 위키피디아 콘텐츠 위주 5억 개의 객체에 대한 Knowledge Graph를 구축, 질의응답 서비스 제공 http://www.google.com/insidesearch/features/search/knowledge.html
AQUAINT Program	NIST*	<ul style="list-style-type: none"> 다국어 기반의 Advanced QA 연구, 테스트, 음성, 이미지, 비디오 지원 http://www-nlpri.nist.gov/projects/aquaint
Project HALO	Vulcal Inc. 2003 ~ 2006	<ul style="list-style-type: none"> Digital Aristotle: 어려운 과학문제 질문에 답변 파일럿 프로젝트(2004년, 6개월): domain expert 개발 AURA(Automated User-Centered Reasoning and Acquisition System) 프로젝트(SRI*, 2004~2006)
Wolfram Alpha	Wolfram Alpha 2009 ~ 현재	<ul style="list-style-type: none"> Siri 서비스의 질의응답 부분 담당, 수식, 알고리즘, 모델을 적용하여 사용자 의도에 맞는 정답을 제시 http://www.wolframalpha.com
Human Brain Project	EU 2013 ~ 2023	<ul style="list-style-type: none"> 인간 뇌의 작동방식에 대한 정확한 이해, 활용을 통해 컴퓨팅 아키텍처, 신경과학, 의학 분야 등에 적용 예정 EU 미래기술 주력 사업 프로그램의 6대 연구과제 선정 http://www.humanbrainproject.eu
Google Brain Project	Google 2008 ~ 현재	<ul style="list-style-type: none"> 9계층의 신경망, 10억개의 연결구조로 인간 인지방식의 자율학습 모델 연구 2만여개의 객체 카테고리를 인식하는데 15.8%의 정확도 http://www.mpi-inf.mpg.de/yago-naga/yago

*NIST(National Institute of Standards and Technology)

*SRI(Stanford Research Institute)

술과 Wolfram Alpha 검색엔진을 이용하여 검색이나 일정 관리 등을 지원하는 Siri 서비스를 개시하였다[6]. Google은 방대한 지식베이스를 이용하여 약 5억개의 지식그래프를 객체를 구축하여 인물, 지역, 사물 위주의 정보를 제공하고 있다.

인간의 뇌를 모사하는 자가학습 기술(self-learning)은 현재의 컴퓨팅 환경의 패러다임을 완전히 바꿀 수 있는 기술이다. 학문적인 관점에서는 자가학습 기술의 역사가 깊은 편이지만 상용화 수준의 기술로 개발하기 위해서는 여러 가지 난제들을 극복해야 하는 고위험 기술 범주에 속하기 때문에 국가 혹은 글로벌 기업이 주도하여 연구가 진행 중이다. Google은 인간의 인지방식을 시뮬레이션하기 위해 약 16,000개의 코어를 이용하여 unlabeled data 기반의 자율학습을 위한 9계층, 10억개의 연결구조를 갖는 신경망에 대한 연구를 진행 중이며, EU에서는 미래 유망 6대 기술 중의 하나로 The

Human Brain 프로젝트를 선정하고 인간의 뇌 동작방식에 대한 보다 정확한 이해와 활용을 통해 컴퓨팅 아키텍처, 신경과학, 의학 등의 분야에서 발전을 도모하고 있다. IBM은 DARPA(Defense Advanced Research Projects Agency)에서 주도하는 SyNAPSE 프로젝트를 통해 256개의 전자 뉴런을 가지는 인간 두뇌를 닮은 컴퓨터 칩을 개발하였으며, 이를 기반으로 향후 7~10년 내에 인간의 지각, 학습, 반응을 모방하는 컴퓨터 시스템의 구축을 목표로 연구 진행 중이다[7].

지식베이스 관련 프로젝트는 주로 대학과 기업에서 진행하고 있다. DBpedia, Freebase, YAGO[8] 등의 위키피디아 기반의 지식베이스들이 구축되고 있으며 이를 이용한 지식추론기술들이 연구되고 있다. CMU에서는 Read the Web 프로젝트를 통해서 자연어를 이해하여 새로운 지식을 학습하고 구축하는 NELL(Never Ending Language Learning) 기술을 연구 중이며[9], Cycorp에

〈표 2〉 지식베이스 관련 프로젝트 현황

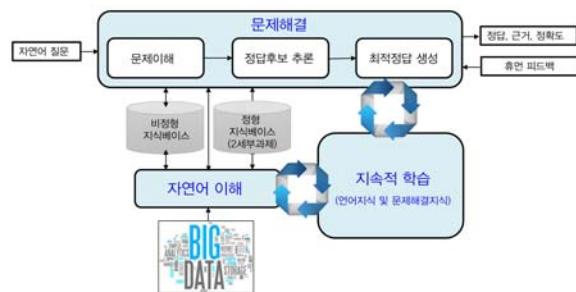
프로젝트명	수행기관 및 기간	주요 내용
NELL	CMU, DARPA, Google, NSF 2010~현재	<ul style="list-style-type: none"> 기계학습 기반 비구조/반구조 웹 데이터에서 구조화된 정보추출 시스템 기반 지식베이스 구축 프로젝트 현재 87%정도의 정확도로 추정되며 818 개 범주 및 관계에 대해서 약 189 만개의 인스턴스로 구성 http://rtw.ml.cmu.edu/rtw
CYC	Cycorp, Inc 1984~현재	<ul style="list-style-type: none"> 세계 최대 규모 지식베이스 및 Common Sense 추론엔진 약 24 만개의 개념 및 2 백만개의 트리플, 6 만 9 천여개 외부 지식베이스 링크 OpenCyc 프로젝트를 통해 지식베이스 공개(6,000 개 이상의 개념 포함) http://www.cyc.com
DBpedia	University of Leipzig Freie Universität Berlin OpenLink Software 2007~현재	<ul style="list-style-type: none"> 위키피디아 콘텐츠로부터 추출된 멀티 도메인 온톨로지 영어버전 DBpedia는 3 백 7 십만개의 개념과 4 억개의 사실을 포함하며, 개체명 위주의 정보를 포함하고 있음 http://www.dbpedia.org
FreeBase	Metaweb Tech. (Google) 2003~2006	<ul style="list-style-type: none"> 위키피디아 콘텐츠로부터 추출된 2 천 3 백만개의 구조화 데이터에 대한 지식 그래프로 구성됨 http://wiki.freebase.com
YAGO	Max-Planck-Institute Saarbrücken 2008~현재	<ul style="list-style-type: none"> 위키피디아, 워드넷, GeoNames 콘텐츠로부터 추출된 1 천만개의 구조화 데이터에 대한 지식베이스 http://www.mpi-inf.mpg.de/yago-naga/yago

서는 세계최대 규모의 지식베이스 및 추론 엔진을 개발하여 OpenCyc 프로젝트를 통하여 공개하였으며[10], IBM은 위키피디아, CIA Fact Book 등 다양한 콘텐츠로부터 구조정보를 추출하고 지식프레임을 구성하여 단답형의 질의응답 기술을 개발하고, 의료, 금융 등의 다양한 분야로 확장하고 있다. 〈표 2〉에는 지식베이스 구축 프로젝트의 현황을 정리하였다.

빅데이터 지식처리 인공지능기술의 국내 연구사례는 그리 많지 않은 편이다. 삼성에서는 S보이스 서비스를 통해 Wolfram Alpha의 단답형 질의응답 서비스와 제휴하여 영어권 서비스를 제공하고 있으며, 다음 포털은 영화, 인물, TV 프로그램 시청률 등의 구조화된 정보를 대상으로 제한된 단답형 질의응답 서비스를 제공하고 있다.

한편, 미래창조과학부에서는 2013년부터 10년간 엑소브레인 SW는 인간의 지식증강서비스를 위해 빅데이터로부터 스스로 학습하여 지식을 축적하고, 시스템 및 기기 간의 자율협업방식으로 새로운 문제를 해결하는 엑소브레인 SW 기술개발 과제를 진행 중이다. 4개의

병렬형 과제로 구성되어 있으며, 지능 진화형의 질의응답, 빅데이터 이해 기반의 자가학습형 지식베이스, 자가학습 지능 원천기술 및 지식협업 프레임워크 기술이 연구 내용이다. 본 문서에서는 지능 진화형 질의응답 시스템, 일명 WiseQA라는 프로젝트에 대해 간략히 소개한다. WiseQA 과제는 엑소브레인 과제 4개의 병렬형 과제 중에서 1세부과제로 개념은 복잡한 자연어로 기술된 문제의 의미를 이해하여 정답을 추론하여 생성하는 것이다. 이를 위해 자연어 질문을 이해하고, 비정형 및 정형 지식베이스를 통해 정답 후보들을 추론하고, 순위화, 필터링 등의 과정을 거쳐 최종 정답을 근거와 함께 추론



〈그림 3〉 WiseQA 과제 개념도



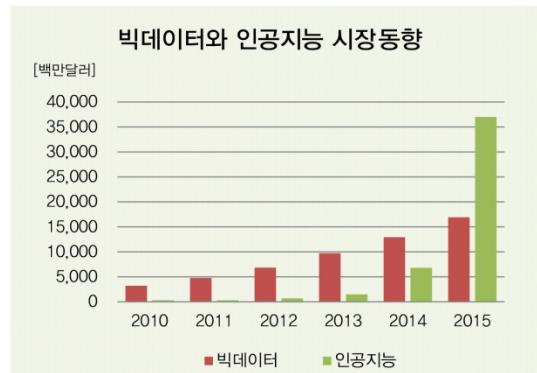
(그림 4) WiseQA 과제의 도전 및 핵심 기술

하는 모듈이 필요하다. 오답에 대해서는 사용자에 의한 피드백이 주어지며 오답과 새로운 지식에 대해서 지속적인 학습을 통해 언어지식과 문제해결 지식을 학습하게 된다. (그림 3)은 WiseQA 기술의 개념도를 나타낸다.

WiseQA 과제에서 도전하게 될 기술은 크게 3가지로 나눌 수 있으며, 자연어의 어휘, 문법, 문맥의 의미를 이해하는 자연어 이해 기술, 빅데이터로부터 끊임없이 새로운 언어지식 및 문제해결 지식을 학습하는 지속적 학습 기술, 자연어 질문에 대한 문제를 이해하고 정형/비정형 지식으로부터 정답을 추론하는 질의응답 기술이 핵심기술이다. (그림 4)는 지능형 소프트웨어의 메가트랜드에 따른 WiseQA 과제의 핵심 및 도전 기술을 나타낸다.

IV. 산업동향

빅데이터 지식처리 인공지능기술과 관련된 시장 상황은 IDC에 따르면 빅데이터 관련 분야는 연평균 40%의 성장률을 보이며 2015년 약 170억 달러로 예상되며, 인공지능 분야는 2013년 8억 달러에서 2015년 370억 달러로 증가되어 빅데이터 분야 시장을 능가할 것으로 예상된다. 이는 인공지능기술이 빅데이터 관련 전분야에서 핵심기술로 등장하게 될 것이기 때문으로 예상되기 때문이다. (그림 5)는 빅데이터 및 인공지능기술의



(그림 5) 빅데이터 및 인공지능기술의 시장동향

〈자료〉: IDC, EU, "Market Report," 2013.

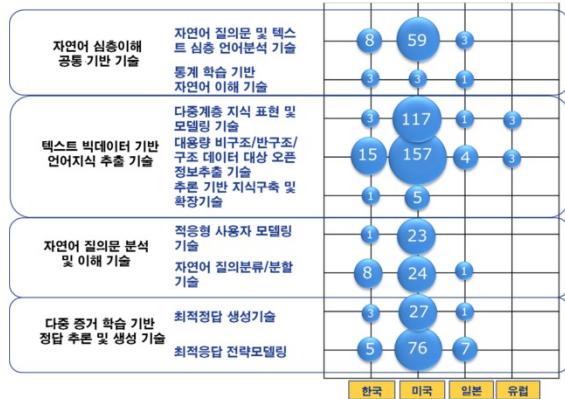
시장 동향을 나타낸다.

(그림 5)와 같이 아직 인공지능 관련 시장은 시장 도입기 단계에 있는 것으로 판단되며, 특히 심층 질의응답 서비스 관련 시장은 현재 기초 기술 확보 단계에 있어 본격적인 상업화 서비스는 이루어지지 않고 있다.

빅데이터 및 인공지능 관련 산업동향을 살펴보면, Google, Microsoft, Yahoo와 같은 검색 엔진 회사에서는 인공지능의 기계학습 기술, 텍스트 마이닝 기술, 자연어 처리 및 이해 기술을 활용하여 지능형 검색 SW의 성능을 향상시키려고 노력하는 중이며, 미국의 DARPA, SRI 등에서는 인공지능 기반 기술인 지능형 계획, 기계학습 기술, 에이전트 기술에 대한 연구를 지속적으로 진행하고 이 연구로부터 인공지능기술의 상업화를 위해 Siri, Twine 같은 회사를 설립하였다.

V. 지식재산권 동향

빅데이터 지식처리 인공지능 SW 기술의 특허현황은 국내의 경우 2003년을 최고점으로, 2000년부터 2010년 까지 등락을 보이고 점진적으로 감소하였으나, 이는 주요 출원인들이 양보다는 질에 치중한 흐름이 반영된 것으로 예상된다. 출원 기업은 Microsoft가 1위, 그 뒤로 IBM과 Google 등이 다수의 특허를 출원하였으며, 미국



(그림 6) 빅데이터 기반 심층 질의응답 기술별 특허맵

에서의 출원이 나머지 전체 국가 출원 규모의 절반 이상을 차지하고 있다. 특히, Google, IBM, Apple 등이 기보유하고 있는 특허를 회피할 수 있는 개량 특허 및 신규 특허의 확보가 시급한 상황이다. (그림 6)에는 빅데이터 기반의 심층 질의응답 기술의 기술별 특허맵을 나타내었다. 자연어 심층이해 공통 기술, 빅데이터 기반 지식 추출 기술, 사용자의 질문의도 분석 및 이해 기술, 정답 추론 및 생성 기술 모두 미국에서 출원된 특허가 다수이며, 한국, 일본 등이 그 뒤를 따르고 있는 상황이다. 기술별로 약간의 유사 특허들이 존재하지만 지능 진화 및 자기학습 방법 등에서 차별화를 통해 국내외 핵심 특허 확보가 가능한 상황으로 분석된다.

제안 가능한 국내외 표준화 그룹 및 내용에 대한 현황

〈표 3〉 주요 기술별 표준화 그룹 및 내용

표준화 단체	표준화 내용
TTA PG606 메타데이터 표준화 그룹	관계형 지식 추출 모델 및 데이터 포맷
ISO MPEG User Description 그룹	사용자별 맞춤형 콘텐츠 표현을 위한 지식 콘텐츠 포맷
ISO TC37 Language and Content Resource 그룹	자연어 의미 프레임 표준
ITU-T SG2 그룹	자연어 질문 형식
ITU-T SG16 그룹	자연어 인터페이스를 위한 대화모델
ISO TC37/SC4 Language Resource Management 그룹	지식 표현체계

은 〈표 3〉에 정리하였다. 앞서 소개한 WiseQA 과제를 통해 ITU-T SG16에 Intelligent QA Service Framework의 신규 표준화 항목이 채택되었다.

VI. 발전방향

IDC에 따르면 현재는 창출되는 데이터의 약 5% 정도만 지식화되고 있으며, 10년 후에는 지금보다 약 50배 가량의 데이터 폭증이 있을 것으로 예상되고 있다. 이처럼 지식화되지 못하고 벼려지는 데이터를 지식화하기 위해서는 데이터의 심층이해 기술과 스스로 학습하고 판단하여 지식을 구축하는 기술, 인간의 뇌를 모사하여 지식 처리가 가능한 인공지능 원천 기술 등이 포함된 지능형 지식처리 SW 플랫폼의 개발이 필수적이다.

이와 관련하여 미국, 일본, 유럽은 물론 국내에서도 국가 차원의 대규모 프로젝트가 추진 중이며, 빅데이터 지식처리 인공지능 SW의 관련 기술이 상용화 수준으로 개발이 된다면 국가 소프트웨어 산업의 미래 경쟁력 확보는 물론 제조업, 도/소매업, 과학 및 교육 분야 등에서도 인공지능 기반의 컴퓨팅 기술 적용이 가능해져 전통산업 ICT(Information & Communication Technology)화가 가속화될 것이다. 또한 기업 및 공공분야의 경영자와 전문가의 정책 및 의사결정 지원, 사회현상 분석과 예측 등의 분야에도 활용될 수 있을 것이다. (그림 7)은 빅데이터 지식처리 인공지능 SW의 산업계 확산 방안 및 기대효과를 나타낸다.

지식처리 소프트웨어의 핵심기술인 지능 진화형 질의응답 기술은 웹이나 모바일에 탑재되어 일반 지식 및 상식에 대해 사용자의 지식을 증강시켜주고 기기나 설비, 빌딩, 의복 등에 장착되어 탑재형의 에이전트로써 사용자의 업무 처리에 효율성을 극대화하고, 교육이나 공공 분야에서는 지능형 튜터로서의 역할을 기대할 수 있을 것이다. 물론 자기학습이 가능한 지식베이스와의 연계가 우선 되어야 할 것이다. 인간의 사고방식을 모사하는



(그림 7) 산업 확산방안 및 기대효과

자기학습이 가능한 인공지능 모델은 정보에 대한 의미 분석 정확도를 향상시켜줄 것이다.

경제적 기대효과를 살펴보면 글로벌 기업 대비 저비용 고효과의 인공지능 기반 지식처리 SW의 산업계화로 산으로 인해 ICT 산업의 부가가치 창출과 의료, 법률, 금융, 기업/공공 정책 등의 문제 해결 및 의사결정 지원으로 인하여 고품질, 저위험의 산업 기반 조성이 가능해져 국가적으로 경제적 소모비용을 절감할 수 있을 것이다. 또한 빅데이터 지식처리 인공지능 SW 요소 기술의 직접 활용 및 타산업 연계 활용과 더 나아가 다국어를 지원하는 글로벌화를 통해 지능형 지식산업분야의 신시장 창출도 가능할 것이다. (그림 8)은 기술의 적용방법 및 산업 연계방법에 대해 나타낸다.



(그림 8) 기술 적용 및 산업 연계 방법

VII. 맷음말

빅데이터 지식처리 인공지능 SW의 핵심기술들과 기술 분야별 국내외 동향에 대해 간략히 살펴보았다. 해외 선진국들은 국가 주도 혹은 글로벌 기업들의 주도하에 지식 혹은 지능처리 소프트웨어에 대한 연구를 진행 중에 있다. 또한 빅데이터에 대한 관심이 증폭되면서 인공지능기술과의 접목이야 말로 데이터의 단순 활용을 넘어 제대로 이해하고 해석하여 고수준의 지식이 생성될 수 있을 것으로 예측되며, 이를 위해서는 우리나라에서 도 관련 분야에 대해 정부의 장기적이고 정책적인 지원이 필요할 것이다.

지식처리 소프트웨어 분야는 현재 해외 의존도가 높은 기술분야이며 우리나라에는 선진국에 비해 경쟁력이 낮은 편이다. 현재 미래창조과학부에서 진행되고 있는 엑소브레인 SW 기술개발 과제를 통해 차별성 있고 도전적인 원천 기술의 개발과 사업 생태계 조성을 통해 웹 및 모바일 플랫폼 이후에 등장할 제 4세대 인공지능 플랫폼의 전쟁에서 승자가 될 수 있기를 기대해본다.

용어해설

엑소브레인 내 몸 바깥에 있는 인공 두뇌라는 뜻

자율학습 비교교 학습으로도 불리는 기계 학습의 일종으로, 입력값에 대한 목표치가 주어지지 않고 데이터가 어떻게 구성되어 있는지를 알아내는 학습방법

지식베이스 전문가 시스템의 구성 요소 중의 하나로, 인공지능이 사용될 분야와 관련된 지적 활동과 경험을 통해서 축적한 전문지식, 문제 해결에 필요한 사실과 규칙이 저장된 데이터베이스

지식증강서비스 빅데이터 분석을 통한 자기학습형 지식베이스를 기반으로 사람, 지식, 기계 간의 상호 이해를 통해 분석, 추론 예측형의 전문가 수준의 지식을 제공하는 지능형 서비스

자율협업 이종 도메인의 지식 및 기기의 확장 적응을 위해 스스로 학습하고 판단 및 협업하는 인공지능기술

약어 정리

AQUAINT Advanced Question Answering for Intelligence

AURA Automated User-Centered Reasoning and
Acquisition System

DARPA	Defense Advanced Research Projects Agency
ICT	Information & Communication Technology
NELL	Never Ending Language Learning
NII	National Institute for Informatics
NIST	National Institute of Standards and Technology
PAL	Personal Assistant that Learns
Siri	Speech Interpretation and Recognition Interface
SRI	Stanford Research Institute

참고문헌

- [1] 성낙환, “인공지능기술의 결음마가 시작되었다,” LGERI 리포트, 2012.
- [2] 김성태, “빅데이터 시대, AI의 새로운 의미와 가치,” IT&Future Strategy, 2012.
- [3] Todai Robot Project, <http://www.fujitsu.com/global/news/pr/archives/month/2012/20120910-01.html>
- [4] DeepQA Project, <http://researchweb.watson.ibm.com/deepqa>
- [5] D. A. Ferrucci, “Introduction to “This is Watson”,” *IBM J. Research Development.*, vol. 56, no. 3.4, 2012, pp. 1:1-1:15.
- [6] Wolfram Alpha, <http://www.wolframalpha.com>
- [7] DARPA, <http://www.darpa.mil>
- [8] YAGO Project, <http://www.mpi-inf.mpg.de/yago-naga/yago>
- [9] Read The Web Project, <http://rtw.ml.cmu.edu/rtw>
- [10] CYC Project, <http://www.cyc.com/platform/opencyc>