



음성인식 기반의 인공지능 서비스 연구 동향

강장목* · 박정호**

*고려대학교 정보대학, **서울디지털대학교

목 차

I. 서론	IV. 음성인식기반의 인공지능 서비스
II. 음성 인식 연구 동향	V. 결론
III. 인공지능 연구 동향	

I. 서론

2016년 구글의 알파고와 이세돌의 대국 이후, 우리나라는 2015년 빅데이터를 넘어 인공지능의 열풍에 휩싸여있다. 알파고 사태로 불을 지핀 후 인공지능 관련 산업의 전망이 좋아지고 해당 사업체의 주식이 상승하고 미래를 기계가 지배할 사회로 단정하거나 두려워하는 기사가 속출하는 등 인공지능에 대한 갑론을박이 대단하다. 강인공지능, 약인공지능, 기계학습 등 인공지능과 연관된 연구주제에 대한 관심이 뜨거운데, 기술의 본질을 제대로 이해하지 못하고 있는 일부 전문가가 흑세무민하여 많은 사람이 인공지능에 대한 막연한 공포심이나 과도한 환상을 갖고 있는 실정이다.[1] 현실의 수준에 대한 냉철한 파악 없이 막연히 딴청과 같은 현재 기술의 선형적인 진보로 SF영화 속의 이야기가 실현될 것처럼 말해서는 곤란하지 않을까?[1]

이 글은 인공지능이 연구 앞단의 유행처럼 접두사로 활용되는 사태를 막고 인공지능기술이 갖는 다양한 응용 가능성을 사례로 보여줌으로 이에 대한 이해를 도모하는 것을 목적으로 한다.

물론 이 글은 인공지능만을 다루는 것이 아니라, 인공지능 기술과 결합하여 라이프로그킹(life logging; 일상생활 현장 속으로 침투하여 삶의 본질적인 영역을 영위하는데 구체적인 서비스 형태로 도움을 주는 기술), 인간 친화적 기술, 이용자 상황 민 문맥 이해 기술 중 음성인식과 결합한 서비스 형태로서의 인공지

능을 살핀다.

음성은 정보 전달방법 중 가장 빠른 수단이며, 가장 친숙한 방법이므로 별도의 훈련이나 학습 없이도 사람들이 로봇을 손쉽게 사용할 수 있다.[2]

심심이, 애플의 시리는 대표적인 대화 처리 기술(텍스트든 음성이든 그 입력이 무슨 의미인지 이해하고 또 반응하도록 하는 것)로부터 발전한 사례이다. 지니톡의 경우에는 음성인식과 언어번역기술 등이 결합된 형태이다. 지니톡은 인공지능기술 중 음성에 기반을 둔 연구의 결과물이다. 인공지능은 구글의 자율주행 차량이나 퀴즈대회에서 승리한 IBM의 왓슨 시스템 등 승패를 따지거나 계산을 하데 현재 활용되고 있다. 이처럼 인공지능과 음성인식은 동떨어진 기술이 아니다. 음성인식 기술의 보다 정교해지고 다양한 도메인에서 활용되기 위해서는 인공지능에 대한 이해와 적용이 반드시 필요하다.

이 글은 음성 인식 기반의 인공지능 서비스 모델의 사례를 발굴하고 이를 분석하여 차세대 인공지능과 음성인식 결합 서비스의 연구 동향을 고찰한다.

II. 음성 인식 연구 동향

음성인식이란 인터페이스(Interface)의 일환이다. 인간과 기계 사이에 의사소통을 음성을 기반으로 한다. 인간과 기계 사이의 의사소통에 음성을 사용하는 것이

타이핑하거나 터치하는 것에 비하여 훨씬 수월하다. 이 수월함은 바로 인간 친화적 기술이라고 표현된다. 예를 들어 저자가 눈을 뜨고 일어나 '물을 한 잔 쥐'라고 말하는 것은 무척 습관적이고 손쉬운 방식이다. 그러면 이 말을 알아들은 가족이 물을 떠주는 경우가 있거나 아니면 편잔을 듣게 된다. 반면, 목소리를 내어서 무엇을 요청하는 이런 자연스러운 행위는 고도의 인간 언어를 이해하는 기계인 경우에 가능하다.

인간과 기계 간의 통신이나 의사 전달 등을 오랜 기간 동안 UI에서 체험 중심의 UX로 발전하여 왔다. 그리고 인터페이스는 인간이 기계와 대화한다는 불편함이나 인식이 없는 자연스러운 환경을 구성한다.

그 중 대표적인 하나가 음성 기반이다. 음성기반이란 입력 받은 음성을 컴퓨터가 분석하고 특징을 추출한 다음 미리 입력된 음성모델 데이터베이스와 비교하여 문자 혹은 명령어로 변환하는 기술을 의미한다.[3-4]

백수지, 이영재(2012)의 연구에 따르면 음성인식은 인간고유의 정보전달 방식으로 별도의 훈련 없이 사용될 수 있는 차별화된 강점을 보유하고 있으므로 자연스러운 인터페이스 기술로 활용될 수 있다. 이동 중이거나 다른 작업을 할 때 음성을 통한 입력이 가능하므로 모바일 기기에 적합한 인터페이스이며 손으로 타이핑하는 것을 대신 할 수 있으므로 장애인이나 노약자에게 효율적인 인터페이스로 사용될 수 있다.[5]

그러나 외부 환경에 소음이 많은 곳에서 사용이 제한된다.[5] 따라서 외부 환경의 소음에 대응하기 위한 여러 알고리즘적 방법론이나 여타 기술적인 접근이 필요하다. 이처럼 음성인식 시스템은 잡음이 들이는 환경에서 사용될 경우를 대비한 연구가 왕성하게 이루어졌다. 노이즈를 제거하거나 조정할 수 있기 위해 잡음에 강인한 특징 추출 방법이나 음성을 개선하거나 음성의 특징과 음향 모델 간의 보상 주는 방법 등이 제안되어 왔다.

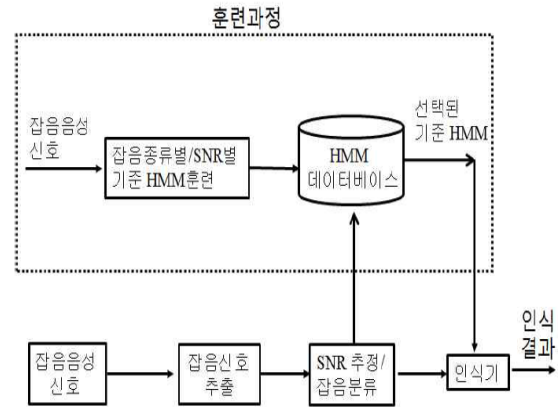


그림 1. 멀티 모델 기반 인식기의 전체 블랙 다이어그램 [6]

출처: 정용주, 음성특징 보상을 이용한 멀티모델 기반 음성인식기의 성능향상, 한국정보기술학회, 한국정보기술학회논문지 10(7), 2012.7, 179-184

위 그림은 멀티 모델 기반으로 음성인식 시스템의 성능을 향상시키기 위한 블랙 다이어그램을 도식화한 것이다. 이 다이어그램을 살펴보면 잡음을 종류별, SNR별 등의 기준으로 구분하여 HMM훈련을 실시한다. 이후 HMM 데이터베이스를 통해 선택된 기준의 HMM만 인식하고 인식기를 통해 인식 결과를 내보낸다.

이상의 과정은 인식된 잡음 음성과 훈련된 잡음 음성 간의 잡음 종류에 대한 차이를 구분해낸다. 그리고 인식잡음 음성과 훈련잡음 음성의 차이에 의한 불일치를 이용하여 인식의 성능 저하를 막아내는 방식이다. 이처럼 음성인식의 성능을 높이기 위한 다양한 아이디어와 노력이 진행되었으나, 최근에는 소셜 네트워크 서비스와 빅데이터 그리고 인공지능 등 음성인식을 하는데 있어 장소적 특징과 음성을 주고받는 객체의 속성을 파악할 수 있는 기술이 등장하여 이들 다양한 기술 간의 조합을 통해 음성 인식률을 높이는 방안도 제시되고 있다.

음성인식의 정확도를 높이는 이상의 방법 외에도 음성인식 시스템은 다양한 응용환경에서 적용되고 활동되면서 그 연구 영역과 실용성을 높여 왔다.

정민화, 이규환, 신대진, 정지오, 강경희(2016)은 음성 인식기술과 응급상황관제를 적용하였다. 음성인식 기술을 응급상황관제시스템에 적용하고자 하는 목적은 상황요원의 119신고접수 시 신속성과 효율성을 지원하고, 상황요원의 올바른 상황판단과 신속한 응급조치를 지원하는데 있고 메뉴얼에 나타난 조치 사항을 상황요원에게 제시하여 상황요원이 적절한 조치를 취할 수 있도록 함으로써 메뉴얼에 따른 시스템과 재량적인 판단에 따른 시스템의 장점을 결합하는 것이라 할 수 있다.[7]

정양근, 조상영, 양준석, 한성현(2015)는 로봇의 교시에 음성인식 기술을 활용한다면 사람에게 보다 친숙하면서 전문가가 아니더라도 로봇의 교시를 가능하게 하여, 전문 엔지니어 양성 및 프로그램 재개발에 따른 비용과 시간의 절감 효과를 가져 올 수 있다고 보았다.[8]

이 외에도 음성인식 연구는 다양한 응용 형태로 발전하고 있다. 예를 들면, 음성인식기술은 최근 자동통번역 외에 스마트폰 덕에 애플 시리가-음성인식 기반의 인공지능 개인 비서 서비스, 구글은 음성인식 기반의 검색 서비스, MS의 음성인식 등이 있다.

이들 서비스의 공통점은 자사가 가진 핵심 경쟁력을 바탕으로 음성 인식 서비스를 발달시키는 것이다. 최근에는 음성인식 기술과 인공지능 기술을 결합한 새로운 형태의 서비스가 소개되고 있다.

III. 인공 지능 연구 동향

인공지능이란 지능을 뜻한다. 인간이 지능을 어떻게 형성하는지도 명확하지 않다. 따라서 ‘인공지능은 A이다.’ 또는 ‘B이다.’라는 식으로 딱히 정의하기 어려운 문제가 있다. 지능이란 것 자체가 명확하게 정의될 수 없는데 인공지능이라고 등장한 기술이나 서비스 역시 그 구현된 형태를 보고 가늠하는 것이지, 엄밀하게 그것 자체가 ‘인공지능의 대표적인 모습이다’라고 단정 지을 수는 없다. 지능을 인공적으로 재현한다는 것이 쉽지 않고 우리 뇌 속에 지능이 구성되고 형성되는 과정 자체도 우리는 명확하게 과학적으로 관찰하고 있지 못한 실정이기 때문이다.

일반적으로 지능은 첫째, 외부를 인식한다.

둘째 지능은 외부를 객관적으로 바라본 후 추론한다. 셋째 지능은 외부 세계와 적응하는 능력이다. 마치 인간이 외부 환경의 추위와 더위 그리고 습기와 건기 등을 스스로 자율적으로 몸이 알아서 조정되어 땀을 흘리고 모공을 수축하는 활동과 같다.

그렇다면, 인간조차 어떻게 지능이라는 기능을 자율적이면서도 능동적으로 수행하는지 명확히 모르는 상태에서 전통적인 환원주의(reductionism)에 입각한 과학적 방법으로는 구현이 어렵기 때문에 인공지능은 인간이 만들어낸 구현된 형태로 어렵잡작하는 수준이 된다.



그림 2. IBM 왓슨의 퀴즈쇼[1]

출처: 조성배, 인공지능 기술전망과 미래산업의 주요과제, 과학기술정책연구원, FUTURE HORIZON 28, 2016.5, 6-9

2011년 2월 TV 퀴즈쇼 <저퍼디(Jeopardy)>에서 우승한 IBM의 인공지능 컴퓨터 ‘왓슨’은 암 진단 분야에서 의사들을 앞지르고 있다.[9] 암 전문의의 초기 오진율은 20~44 퍼센트에 달하는데, 왓슨은 2~9 퍼센트밖에 안 된다. 크레디트 스위스 은행은 인공지능을 사용해 투자 의사결정을 돕는 보고서를 내고 있는데, 투자 분석가보다 세 배 더 많은 양을 써내면서도 질이 높다고 한다. 구글의 인공지능 개발책임자 레이 커즈와일은 2029년이 되면 인공지능이 모든 면에서 인간을 앞설 것이라 공언한다.[10]

박병원(2016)은 18세기 산업혁명 당시 증기기관, 제니방직기 등의 사례를 시작으로 철도, 전기, 내연기관, 인터넷 등의 발명으로 인한 사회경제시스템의 극적인 변화를 이미 여러 차례 경험하였고 ‘인공지능’ 로봇, 빅데이터를 활용한 소위 말하는 ‘제4차 산업혁명’ 또한

그 중에 하나가 될 가능성이 높다고 보았다.[10]

조성배(2016)은 1956년 다투머스 회의에서 인공지능이란 용어가 만들어진 이후로 먼저 시도된 방법은 인식, 추론, 학습과 같은 지적 기능을 모방하기 위해 이를 보유하고 있는 사람이 해당 영역의 지식을 기호로 표현하여 저장하고, 이를 논리적인 규칙에 입각하여 처리함으로써 문제를 지적으로 해결하고자 시도한 전자의 방법이었다고 말한다.[1] 지금도 전문가 시스템이나 논리/탐색기반 문제해결 방법과 같은 형식으로 널리 사용되고 있으나 반면에 데이터기반 방법론은 기계 학습이나 데이터마이닝이란 이름으로 각광받고 있는데, 해당 문제의 사례를 데이터로 제공하고 이로부터 연역적으로 지식을 추출하여 문제를 해결하는 것이다. 데이터로부터 연역적으로 모형을 구축하는 것은 오래 전부터 전통적인 통계나 확률로 시도하던 방식이기 때문에 기계학습의 많은 방법들이 이에 기반 한 것이 많으나, 이러한 방법의 제약을 극복하기 위하여 신경망(neural network)과 같은 다소 융통성 있는 방법이 시도되고 있다.[1]

이상의 연구 동향에서처럼 인공지능은 오래전부터 인간의 지성을 흉내 내는 연구로 발전하다가 최근 빅데이터 등 다양한 비정형데이터 분석 기술과 함께 더욱 각광받고 있다. 특히, 인공지능이 음성인식 서비스와 결합하여 새로운 비즈니스 모델을 선보이는데, 이를 살펴보면 다음 장과 같다.

IV. 음성인식기반의 인공지능 서비스

음성인식 기술 기반의 인공지능 서비스는 하이브리드한 형태의 서비스 개발이다. 이 글에서는 음성인식 기술을 분석하고 인공지능의 여러 응용 사례를 본 후, 본 장에서 음성인식기반의 인공지능 서비스의 구현 과정과 방법 그리고 문제점 등을 고찰한다.

따라서 이 장에서는 음성인식 기반의 인공지능 서비스가 갖는 가치를 평가한다. 그 가치란 세상에 없던 새로운 정신 즉 공유 서비스, 개인 및 맞춤형서비스, 비대면 서비스로 나뉜다. 이 비즈니스적인 가치의 전환 과정을 우선 다루고 두 번째로 음성 인식 기술 기반의 인공지능의 융복합 형태인 서비스 구현의 여러 고려 사항을 고찰하고자 한다.

첫째, 인공지능은 기계학습과 신경망 이론만으로 현실 세계에 파급효과를 주지는 않는다. 오히려 인공지능은 현실 세계에 등장한 다양한 기술을 응용하는 과정에서 그 파급력을 높이기 때문이다.

인공지능은 전통적인 비즈니스 모델에서 “자원/활동-가치제안-관계/채널” 등의 구성을 달리하거나 전에 없던 새로운 가치로 재편한다.

예를 들면 인공지능이 구글 자동차와 결합하면 단순히 자율운전시스템에 그친다. 사고가 적고 혼자 운전하는 자동차 그 자체가 갖는 시장 가치는 크지만, 이를 이용한 새로운 서비스 모델의 파급효과는 자승으로 크다. 오늘날 ICT 혁명의 새로운 서비스인 공유경제 모델 우버의 경우 인공지능을 이용한 자율주행차량과 결합하는 서비스 모델을 탐색하고 있다. 우버는 택시 운전사, 자가용운전사를 대신하는 인공지능 기반에 공유경제 모델을 접목하여 차량을 가진 누구나 운전 대행 서비스 또는 운송 서비스를 하도록 하는 전에 없던 시장을 개척한다.

따라서 인공지능은 그 자체 기술로 머무르지 말고 새로운 비즈니스 환경과 맞닥트려 전에 없던 가치를 생산해 내는 방향으로 발전해야 할 것이다.

기존의 비즈니스 모델의 가치 사슬이 신기술로 변화하는 모습을 간략히 담은 그림은 아래와 같다.



자료: 이성호·설라영·김은희(2015), 『신기술 발전에 따른 산업 지형의 변화 전망과 대응 전략: 제 2권 인지컴퓨팅』, 정책연구 2015-12-02, 과학기술정책연구원.

그림 3. 인공지능이 견인하는 비즈니스 모델 변화 [11]

음성인식 기술이 인공지능과 만나면 애플의 슈리와 같은 개인 맞춤 비서 서비스가 가능하다. 위 그림에서 기존의 가치제안이 개인맞춤서비스로 볼 수 있다. 또한 음성인식 기술이 구글 자동차를 만나면 목적지와 가는 길 등을 음성으로 묻고 이에 대한 최적의 경로를 구글 기반의 nawro로 찾아주는 서비스가 가능하다.

이는 위 그림에서 관계/채널이던 과거 가치가 비대면비스로 옮겨진 예이다. 또한 인공지능과 우버가 결합한 경우에는 기존의 '자원/활동'적 가치가 '공유서비스'로 전환된 것을 발견할 수 있다. 이처럼 인공지능과 음성인식 등의 최신 기술은 각자 발전하기보다 상호 보완하면서 성장하고 있다.

둘째, 하이브리드 인공지능 기술이 보여주는 음성인식 기술 기반의 인공지능이 구성하는 세부 구현 과정에서의 연관 기술(기계학습, 지각인식, 지식 추론, 빅데이터, 사물인터넷)과 효율(효율성, 편의성, 생산성) 등을 고찰하고자 한다.

근래 들어 성공적인 인공지능 시스템을 분석하면 다음과 같다. 성공한 인공지능은 앞서 소개한 어떤 한 기술만을 사용했다기 보다는 문제의 해결방안을 구조화한 후 여러 가지 기술을 조합하는 능력이 뛰어났다고 볼 수 있다.

여러 가지 기술과 서비스를 적재적소에 복합적으로 활용하는 큐레이션 기반의 S/W 구조를 구성하는 방식이 일반적이다.

이 글에서도 음성인식이라는 상당히 진전되고 음성인식률도 높은 여러 기술을 차용하여 인공지능과 결합하면 보다 인간친화적 또는 사용자 중심의 환경을 문맥으로 정확하게 인식하는 서비스가 가능하지 않을까라는 의문에서 시작했다.

예를 들면 최근 인기 검색어가 된 알파고의 경우에도 전체 구조는 탐색 기술을 따른다. 그러나 이 한 가지 전체 기술인 탐색만 의지하지 않고 세부적으로는 탐색의 가짓수를 줄이고자 신경 네트워크 기술을 사용하였다.

또한 IBM 왓슨은 퀴즈 풀이에서 전 세계 챔피언을 이긴 이후 꾸준하게 발전하여 왔다. 최근에는 의사보다 정확하게 암 등을 진단할 수 있는 인공지능으로 발전하여 오진율을 줄이고 있다.

IBM의 왓슨의 경우에도 상식수준의 다양한 지식을 체계적으로 표현하는 규칙기반 시스템을 기반으로 한다. IBM의 왓슨은 정답에 대한 가설을 세운다. 그리고 이 가설을 체계적으로 줄이는 과정에서 기계학습

방법을 사용한다. 즉 IBM의 왓슨은 수백 가지의 인공지능 알고리즘을 어느 한 알고리즘에만 절대적으로 의존하지 않고 복합적으로 사용하여 정확도를 높였다.

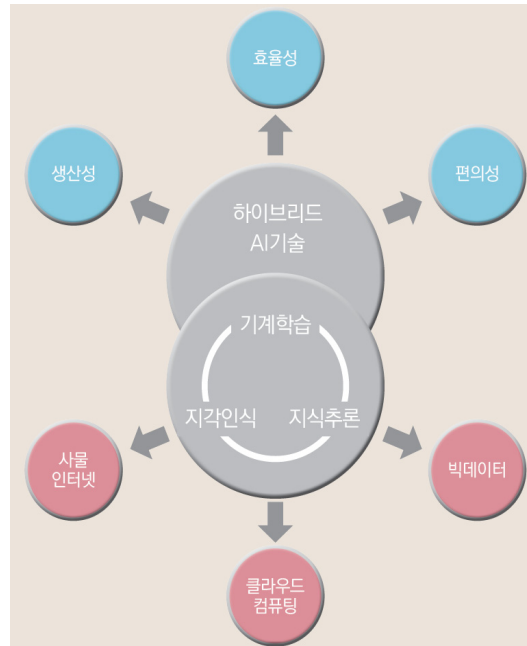


그림 4. 산업적 응용을 위한 하이브리드 인공지능 기술의 구성 요소와 지원구조[1]

출처: 조성배, 인공지능 기술전망과 미래산업의 주요과제, 과학기술정책연구원, FUTURE HORIZON 28, 2016.5, 6-9

여러 알고리즘과 동시에 여러 서비스 기술 등에 대한 장/단점을 정확히 이해하는 것이 중요하다. 위 그림은 빅데이터, 사물인터넷, 클라우드 컴퓨팅이 하이브리드 인공지능과 어떤 관계를 맺고 발전 중인지도 도식화한 것이다. 물론 위 그림은 기계학습, 지각인식, 지식추론 등의 영역에서 각각 발전 중인 세부 기술이 하이브리드 인공지능과 결합하여 편의성, 효율성, 생산성을 극대화하는 것을 잘 보여준다.

최근 IDC, 트랙티카, 맥킨지, 지멘스 등은 세계 인공지능 시장이 급속 도로 증가할 것이다. 사물 인터넷으로 센싱 데이터가 올라오는 경우에는 더욱 더 많은 인공지능 서비스가 요구된다. 이때 인공지능이란 교육 환경에서 올라온 책상 정보, 결상 정보, 칠판 센싱 정보, 교실의 CCTV 센싱 값 등을 결합하여 처리할 수 있는 학습 관련 인공지능이 될 것이다.

즉 인공지능이란 도로, 물류, 교육, 의료, 금융 등 각 도메인에서 최적화된 상태로 발전할 것이다.

인공지능 역시 가장 왕성한 융복합 서비스가 될 전망이다.

V. 결 론

음성인식 기술은 이미 상당히 진전된 기술이다. 기업 입장에서는 음성인식 기술의 디테일을 살려 차별화를 꾀할 것이다.

구글은 검색 기반의 음성 인식으로 애플은 개인화된 서비스인 슈리 기반의 비서 서비스로 차별화되고 있다.

우리나라의 글로벌 서비스인 네이버의 라인도 동남아시아 등의 나라와 우리나라 또는 영어와 여타 국가간의 언어 번역 등의 서비스 기반의 음성인식을 시도하면 어떨까 싶다.

또한 카카오는 전화번호 기반의 관계형 음성인식 서비스를 구현할 수 있다.

이상의 네이버와 카카오의 다음 서비스는 기존 서비스의 핵심 경쟁력 바탕위에서 인공지능과 음성인식이 결합된 서비스로 진화하는 것이 기술 습득과 활용의 수월성에서 바람직하다고 여겨 제안한다.

이처럼 음성인식은 인공지능 기술을 통해 세부 도메인 분야에서 고도화시킬 수 있다. 반면, 기술은 여전히 우리나라 현장에서 적극적으로 도입되고 활용되기에 지나치게 많은 규제와 문제로 인해 그 폐해가 심각하다. 이미 우리나라는 규제당국의 잘못된 여러 처방으로 인해 우리나라만 유독 전 세계 기술발전과 궤를 달리하는 방식으로 성장하는 갈라파고스적 ICT 강국으로 변질되었다.

마지막으로 과거 통상 인공지능과 음성인식 등이 포함된 신기술(뉴미디어)이 문제를 일으켰을 때, 정부당국과 사업자의 패턴을 잠시 기술하며 이 길을 걷지 말기를 조언하고 싶다.

정부는 단기간에 안전한 기술 활용과 신뢰할 수 있는 콘텐츠 제작과 활용에 애쓴다는 홍보효과를 얻기 위해 애쓴다. 그 결과, 공청회, 전문가 간담회, 가이드라인, 지침 등을 쏟아낸다.

대부분의 경우 그 초창점은 시청률 또는 시청자 몇 %이상의 사업자에게는 의무적으로 모니터링 요원을 수십 명 이상 두게 한다는 규범 또는 공인인증서를 받

드시 설치해야 한다거나 일정 수준 이상의 보안 모듈을 탑재해야한다는 의무 조항의 규제 방침이었다.

정부의 발 빠른 규제 대응은 이를 통해 국민을 안심시키고 자신의 역할을 다했다고 인식을 주지만, 선부른 규제 정책으로 장기적인 차원에서 국가 발전 모형이 뒤흔어지는 경우가 다반사였다.

과거 미디어 다음과 네이버 등의 포털을 규제할 때 역시 비정규직 모니터링 요원의 채용을 통해 악플과 댓글의 책임성을 높이려는 시도가 규제 정책 처음에는 어느 정도 이 문제를 해결하는데 기여하였다.

이런 국가의 조치에 사업자의 반발도 만만치 않았다. 그 결과 사업자는 경쟁적으로 필터링 기술을 개발하거나 신사업에 영감을 줄 평판 시스템을 글로벌 수준으로 정교화 하는 노력을 하지 않게 된다.

정부의 규제대로 모니터링 요원을 두거나 기술적으로 댓글을 달 때 실명제를 하도록 한다거나 하는 규제만 지키면, 어느 정도 기업이 저야할 책임으로부터 자유로워지기 때문이다.

이미 글로벌 기업 아마존은 평판 시스템으로 자정 기능을 갖는다. 인공지능의 위협 또는 기계사회가 물고 울 직장을 잃을 사람들의 불안 등으로 규제 당국이 새로운 규제 정책을 낸다면 어떻게 될까? 과연 10년 대계 아이티 산업을 보호하고 키우는 결과가 될지 지금 위에 기술한 다른 IT 영역처럼 오히려 문제를 잠시 막는 역할을 할지 의문시 된다.

우리나라 근성을 솔뚜껍이라고 비아냥거리곤 한다. 문제가 생기면 이를 깊이 있게 성찰하고 긴 안목에서 사업자와 시민의 발전을 도모하기 보단, 들끓는 비난을 면피하기 바쁘다.[12] 그 결과 정부는 19세기적 규제를 제시하고 기업은 그 규제만 지키면 어느 정도 책임을 면피한다. 새로운 미디어가 문제의 심각성을 내뿜을 때 마다, 영혼 없는 공무원과 학·석·박을 오로지 방송과 통신 한 영역에서만 바라본 매몰된 전문가, 그리고 여론의 들끓는 비난을 면피하기 바쁜 사업자 간의 양철 냄비보다 가벼운 규제 모델을 내놓는다.[12]

음성인식 기반의 인공지능 서비스 역시 여타 기술을 조망하고 그 기술을 큐레이션 방식으로 잘 조합 후 세부 도메인에 최적인 서비스를 구현해 내는 역할이 중요하다. 이 과정에서 발생할 주요 사회적 현안을 미리 예단하고 어느 정도 그 방향성이라도 가늠하는 연구 노력이 동시에 필요하겠다.

이 글은 음성인식과 인공지능 기술의 연구 동향과 이 둘을 합친 새로운 서비스의 가치와 비즈니스 모델을 분석하여 향후 음성인식과 인공지능 간의 결합 서비스 제안에 도움을 줄 것으로 기대된다.

Acknowledgement

본 논문은 고려대학교에서 지원된 연구비로 수행되었음(Supported by a Korea University Grant)

참고문헌

- [1] 조성배, 인공지능 기술전망과 미래 산업의 주요과제, 과학기술정책연구원, FUTURE HORIZON 28, 2016.5, 6-9
- [2] H. J. M. Steeneken and D. A. van Leenwen, "MultiLingual Assessment of speaker independent large vocabulary speech-recognition systems: SQUALE Project," Proc. EUROSPEECH-95, pp. 1271-1274 September 1995.
- [3] 김동민, 이칠우, "스마트폰 사용자 인터페이스 기술 동향", pp15-26, 2010.05. 정보과학회지 특집원고.
- [4] 이윤근, 음성인터페이스 기술개요 및 스마트폰 환경에서의 서비스 동향, pp3-9, 스마트폰을 위한 A/V 신호처리 기술, APRIL, 2012
- [5] 백수지, 이영재, 음성인식에 기반한 스마트폰 게임 인터페이스, 한국정보기술학회, Proceedings of KIIT Summer Conference, 2012.5, 454-458
- [6] 정용주, 음성특징 보상을 이용한 멀티모델 기반 음성인식기의 성능향상, 한국정보기술학회, 한국정보기술학회논문지 10(7), 2012.7, 179-184
- [7] 정민화, 이규환, 신대진, 정지오, 강경희, "음성인식 기술의 응급상황관제 적용," 한국화재소방학회, 한국화재소방학회 학술대회 논문집, 2016.4, 77-78
- [8] 정양근, 조상영, 양준석, 한성현, "화자독립방식에 의한 음성인식 알고리즘 개발 및 실시간 실현에 관한 연구," 한국산업융용학회, 한국산업융용학회 논문집 18(4), 2015.11, 250-258
- [9] 유정식, "알파고와 인공지능", 샘터사, 월간 샘터 555, 2016, 50.
- [10] 박병원, 인공지능, 로봇, 빅데이터와 제4차 산업혁명, 과학기술정책연구원, FUTURE HORIZON 28, 2016.5, 4-5
- [11] 이성호, 인공지능 기반의 비즈니스모델 혁신, 과학기술정책연구원, FUTURE HORIZON 28, 2016.5, 10-13
- [12] 강장목, 인터넷 방송 플랫폼 서비스 규제 방안에 대한 고찰(사업자/기업/BJ의 사회적 책무 등 현황 및 개선 방안), KISO저널, 23호 기획동향, 연결: <http://journal.kiso.or.kr/?p=7502>



강장목(Jang-Mook Kang)

고려대학교 정보보호대학원 공학박사(정보보호 전공)
 현) JM 코드 그룹 대표
 현) 고려대학교 정보창의교육연구소 교수
 현) 함께하는 시민행동 운영위원
 현) 경실련 소비자주권국 위원
 현) 국가 인권위 정보인권국 자문 위원
 ※관심분야: 프라이버시, 개인정보, 빅데이터, 정보보호, 소프트웨어 개발 방법론



박정호(Jung-ho Park)

1987년 2월: 성균관대학교대학원 전자공학과(통신공학석사)
 1998년 8월: 성균관대학교대학원 전자공학과(통신공학박사)
 1987년 1월 ~ 1992년 5월: 삼성종합기술원 전자기기연구소 주임연구원
 2004년 3월 ~ 현재: 서울디지털대학교 소프트웨어융합학과 부교수
 ※관심분야: 빅데이터, 지적재산권 및 정보보호, 이터닝, 인공지능, 사물인터넷