

KAKAO

Vol.11

AI

2018.03

AI CODE

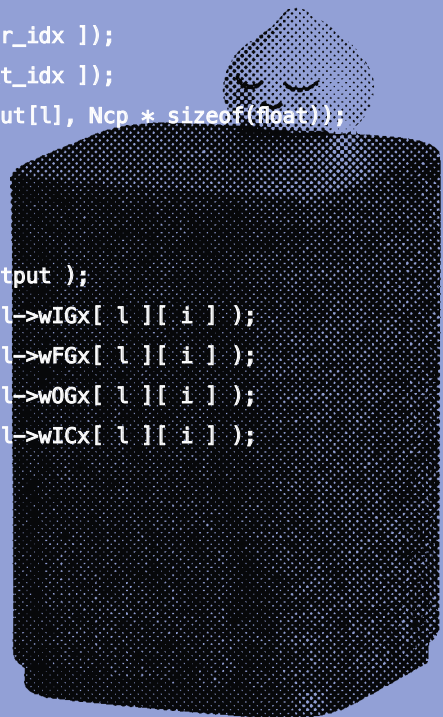
```
static void lstm_calc_outprob_f( float *feat_vec, DNNLSTM *acmodel, float *s_cache, int nfeat , int feat_dim, float **mm_LOutput, float **mm_pt_out, float **mm_pt_inc )
```

```
{
```

```
    ...
    for( t = 0; t < nfeat; t++){
        // each layer input_data : previous_layer_output + previous_time_output
        curr_idx = 0;
        next_idx = num_max_node32;
        // 2. 2nd_layer ~ (last-1)_layer
        Ni = acmodel->num_node[ 0 ];
        curr_output = &(mm_LOutput[ t ][ curr_idx ] );
        memcpy( curr_output, &(feat_vec[ t * feat_dim ] ), Ni * sizeof(float));
        for( l = 1; l < num_layer - 1; l++){
            Nc = acmodel->num_node[ l ];
            Ni = acmodel->num_node[ l - 1 ];
            Nip = Ni * acmodel->prj_rate[ l - 1 ];
            Ncp = Nc * acmodel->prj_rate[ l ];
            Np = Ncp + Nip;
            curr_output = &(mm_LOutput[ t ][ curr_idx ] );
            next_output = &(mm_LOutput[ t ][ next_idx ] );
            memcpy( &(curr_output[Nip]), mm_pt_out[l], Ncp * sizeof(float));
            // input_gate
            pt_inc_ptr = mm_pt_inc[ l ];
            for( i = 0; i < Nc; i++){
                out_avx_l = (v4sf *) (curr_output );
                wgt_avx_ig = (v4sf *) (acmodel->wIGx[ l ][ i ] );
                wgt_avx_fg = (v4sf *) (acmodel->wFGx[ l ][ i ] );
                wgt_avx_og = (v4sf *) (acmodel->wOGx[ l ][ i ] );
                wgt_avx_ic = (v4sf *) (acmodel->wICx[ l ][ i ] );
```

```
    ...
```

kakao



KAKAO AI REPORT

Vol. 1.11

발행일 2018년 03월 30일
발행처 (주)카카오
발행인 여민수, 조수용
편집인 김대원
편집 강연주, 문정빈, 양현서, 윤위훈, 이수경, 인치원, 정하늬, 최은필
디자인 허진아
메일 kakaoireport@kakaocorp.com
브런치 주소 https://brunch.co.kr/magazine/kakaoireport

본 카카오 시리포트의 판권은 주식회사 카카오가 소유하고 있으며, 카카오는 카카오 시리포트를 상업적 목적으로 이용하는 것을 금지합니다. 또한 저작권법에 따른 정당한 범위 내에서 카카오 시리포트를 이용·가공·인용을 하시는 경우라도 반드시 출처를 밝혀주시기 바랍니다.

COVER

카카오 시리포트의 표지에선 AI와 관련된 의미 있는 코드들을 매월 소개하고 있습니다.

Vol.11 코드 | 정대성 denzeljung@kakaocorp.com

카카오미니의 wake up 엔진에서 사용하는 음향 모델 확률 값 연산 알고리즘 일부를 발췌했습니다. Wake up 엔진은 잡음 환경에서도 음성 인식 성능이 우수한 neural network 알고리즘을 적용하여 기존의 방식인 가우시안 혼합 모델(Gaussian mixture model, GMM) 대비 큰 폭의 성능 개선을 이루었습니다.

contents

preface	02
AI & Environment	
AI로 미세먼지와 맞서다	
김동식 AI를 활용한 미세먼지 측정	06
권순박 지하철 내 미세먼지와의 싸움, 그리고 AI	10
AI & Medical	
의료 AI의 역사 그리고 발전 과제 part.1	
정세영 의료 AI의 안착 과정	16
유소영 헬스케어 빅데이터 딜레마와 해결 방안	24
Kakao Mini	
카카오미니에게 궁금한 것	
정대성 박종세 “헤이, 카카오!”를 불러야 하는 이유	32
김다현 카카오미니는 어떻게 점점 더 똑똑해지는가	38
Kakao Brain section	
카카오브레인의 1년 간의 발자취	
이수경 인치원 카카오브레인의 1년, 그리고 미래	44
이수경 안다비 음성은 미래의 인터페이스다	50
이수경 고독의 시대와 AI	54
information	
카카오 AI 전문가 모집 & 연구 지원 프로그램	60
closing	62

카카오 시리포트 11호를 내며

어느덧 추웠던 겨울이 지나가고, 봄이 우리들 앞에 성큼 고개를 내밀었습니다. 지난겨울이 유난히 춥고 길어서인지, 지금 마주하고 있는 봄이 더욱 반갑기만 합니다.

봄이 오는 건 반가운 일이지만, 함께 오는 미세먼지를 생각하면 마음이 벌써부터 답답합니다. 몇 년 전부터는 따뜻한 바람과 함께 전국을 뒤덮는 미세먼지도 우리가 생활하는 공간 곳곳을 찾아옵니다.

걱정이 인공지능(artificial intelligence, AI)과 맞닿자 호기심으로 연결됐습니다. ‘AI가 미세먼지 해결에는 어떤 도움을 주고 있을까?’ 그래서 이번 호에는 AI와 미세먼지만 두 키워드를 모두 다룰 수 있는 전문가 두 명을 섭외하였습니다. 기상예보와 미세먼지 예측에 적용된 AI 기술의 사례와 지하철 미세먼지를 효율적으로 관리하고 개선하는 방법을 담았습니다.

AI가 현장에 적용되어 유익한 역할을 하는 분야 중 하나가 의료입니다. 카카오 시리포트는 작년 7월에 6명의 전문가를 섭외하여 의료와 AI 융합의 모습을 다각도로 분석한 바 있습니다. 그래도, 부족한 부분이 남아 있다고 생각해서 두 글을 추가로 이번 호에 담았습니다. 의료 진단을 소재로 의료분야로의 AI의 적용 역사를 훑었고, 빅데이터와 의료 플랫폼의 원활한 결합을 위한 필요 과제를 정리했습니다.

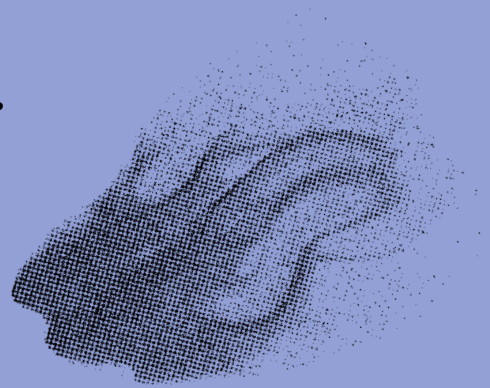
‘AI in Kakao’에서는 카카ومی니에 주목했습니다. 카카ومی니를 깨우는 호출 명령어 “헤이, 카카오!”를 인식하는 과정과 음성 인식 AI 기술의 설명, 그리고 카카ومی니의 진화 양상에 대한 소개 글을 이번 호에 담았습니다.

2017년 2월 세상 밖으로 처음 모습을 드러낸 카카오 인공지능 연구소 ‘카카오브레인’이 쉬지 않고 달려온 지난 1년 간의 발자취, 카카오브레인을 지탱하는 힘이자 카카오브레인을 움직이는 원동력인 3대 연구 원칙, 그리고 앞으로의 추진 방향에 대해 소개하려 합니다. Brain’s pick에서는 음성변환 AI 기술과 미래 인터페이스 발전 방향을 알아보고, AI in pop-culture 세션에서는 초연결사회를 살아가지만 그 어느 때 보다 고독한 시대를 살아가고 있는 현대인들의 모습과 우리가 당면한 과제들을 소개합니다.

이번 호 발간을 위해 좋은 글들을 전해주시신 분들께 다시 한번 감사 인사를 드립니다. 담긴 콘텐츠가 많은 분께 도움이 되고 사회 여러 분야에서 활용되기를 희망합니다.

2018년 03월 30일
카카오 정책지원파트 드림

AI로 미세먼지와 맞서다



살을 에는 듯한 추위로 우리를 힘들게 하던 겨울이 가고, 봄이 왔습니다. 따뜻해서 좋긴 한데, 미세먼지를 생각하면 답답하기만 합니다. ‘골칫거리인 미세먼지를 대비하는 데 AI가 무언가 기여를 하고 있진 않을까?’라는 생각에서 두 전문가를 섭외하여 글을 받았습니다. 두 글에는 대기와 일상의 공간, 그리고 지하철의 미세먼지 관리에 AI가 활용되고 있는 현황을 자세히 설명하였습니다.

AI를 활용한 미세먼지 측정

“기후변화 대응 실패는 단 한번으로도 전 세계에 막대한 영향을 끼칠 수 있습니다.” 2016년 다보스포럼에서 발표한 상위 10대 글로벌 리스크 중 단 한번의 실패로도 가장 큰 영향을 줄 수 있는 것으로 기후변화 대응 실패가 꼽혔다. 이듬 해인 2017년 다보스포럼에서도 기후변화 대응 문제는 가장 시급한 화두로 꼽혔다. 그렇다면 기후변화에 어떻게 대응해야 할까? 유엔(UN)의 기후변화 정부간위원회(Intergovernmental Panel on Climate Change, IPCC)는 기후변화 대응에 대한 해결 방안으로 인공지능(AI), 빅데이터, 사물인터넷(Internet of Things, IoT) 기술을 제시했다. 이러한 기술은 기후변화 대응의 기초라 할 수 있는 기상 예보의 정확도를 높여 보다 정밀한 기후변화를 예측하게 했다. 미세먼지에 대한 정확한 예보를 통해 앞서 언급된 기술이 기상 예보에 적용될 때 어떤 효과를 보여줄 수 있는지를 확인할 수 있다.

글 | 김동식 kdsik@kweather.co.kr

서울에서 태어나 현대 고등학교를 졸업했다. 한양대학교를 수석으로 졸업하고 MIT 기계공학과 석사를 마쳤다. 박사 과정을 밟다가 경영자의 길로 전환하여 세계적인 컨설팅그룹 ADL(Arthur D. Little)에 입사해 경영 컨설턴트로 근무했다. 1997년 민간예보사업제도가 처음 도입되자, 우리나라 최초의 기상 서비스기업 케이웨더를 설립했다. 국내 1위 기상 서비스기업 케이웨더 대표로서 '날씨 경영 전도사'라고 불릴 만큼 기상 산업을 위해 많은 활동을 해왔다. 지금은 한국기상협회 이사(1997~현재)와 기상산업연합회 회장(2002~현재)을 역임하고 있다. 저서로 '날씨 읽어주는 CEO', '미세먼지 극복하기' 등이 있다.

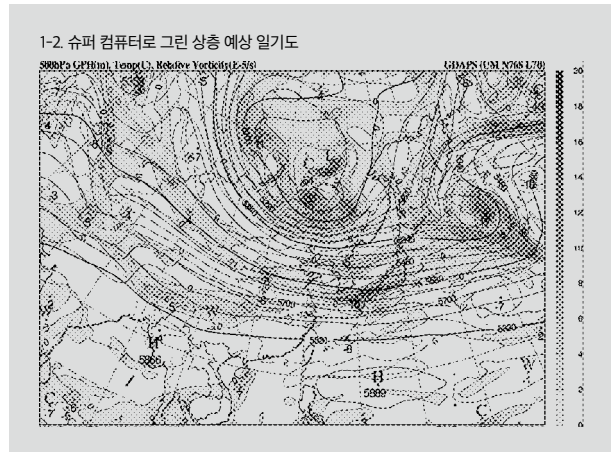
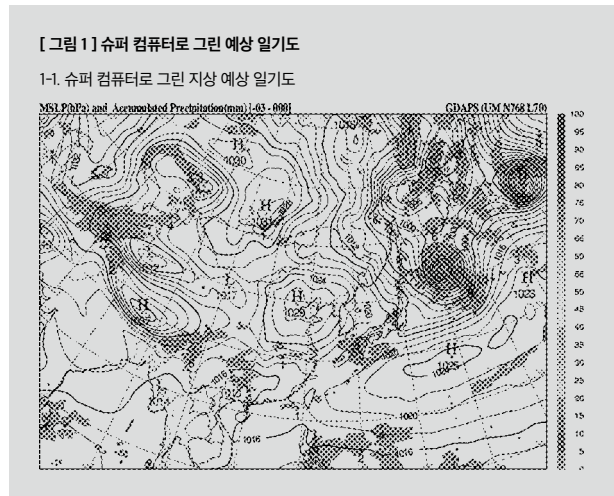
4차 산업혁명 기술 적용 이전의 기상 예측

기상 예보의 생산 과정은 크게 관측, 자료 처리, 현재 일기도 작성, 예상 일기도 작성, 통보의 순서로 이루어진다. 먼저 지표면 근처에서 관측할 수 있는 공기의 온도, 습도, 기압, 바람의 방향 및 속도, 구름의 형태 및 양, 황사나 안개 등을 관측한다. 지표면 근처의 관측은 37개소의 공식적인 기상 관측소에서만 이뤄진다. 500대의 무인자동기상관측장비 자료는 참고로 활용된다. 전국에 8곳이 있는 고층기상관측소에서 하늘 높은 곳의 공기 움직임을 측정한다. 이렇게 수집된 기상 관측 자료들은 기상청의 중앙서버로 취합된다.

표준화 및 보정 작업을 마친 관측 자료와 세계기상기구를 통해 받은 외국 관측 자료를 이용해 슈퍼 컴퓨터로 현재 기상 상태를 나타내는 일기도를 작성한다. 이 때 작성된 일기도는 표준 등압면 일기도와 여러 기상 요소를 볼 수 있는 보조 일기도들이다. 그 후 일기도와 함께 기상위성 자료, 레이더 자료 등을 취합해 수치 모델을 이용한 미래 예측 일기도를 만든다. 예보관들은 수치 예측 모델 자료를 참고하여 날씨 예보를 하게 된다. 결정된 예보는 예보문으로 작성돼 언론 기관이나 인터넷, 유선 등을 통해 일반에게 제공된다.

현재까지 우리나라에서 수행하는 기상 정보 제공 프로세스에서는 AI나 빅데이터를 활용하지 않고 있다. 관측에서 일기도 작성, 미래 예측, 생산된 예보 전파까지 4차 산업혁명 기술과는 동떨어져 있는 것이다. 이는 기상 예측의 낮은 정확도와 낮은 속도, 기상 정보 서비스의 다양성 부족으로 연결된다.

이를 해결하기 위해서 기상청에서도 4차 산업혁명에 관련된 연구를 시작했다. AI 날씨 예보 연구회를 활성화하고 드론 활용 기술 테스트를 진행 중이다. 기상관측 자료에 대한 정확성 및 관측 조밀성 확보를 위해 IoT를 활용하여 온도, 습도, 조도 등의 정보를 수집한다. 이외에 컴퓨터 그래픽과 같은 디지털 기술을 활용하여 기존의 날씨 콘텐츠를 표현의 한계를 극복한 실감형 콘텐츠로 발전시킨 새로운 부가가치를 만들어 내고 있다.



“단기 예보가 정확해지기 위해서는 관측 정확도 향상과 관측소 증가가 필수적입니다. 그러나 인력과 예산 등의 문제로 기상관측소 확장은 매우 어렵습니다.” 전 기상청장의 말이다. 이에 대한 해결책으로 그는 IoT를 이용한 관측자료 확장 방법을 말한다. 서울에서 기온, 강수를 실시간으로 측정할 수 있는 기상관측소는 30개소 뿐이다. 이 정도의 관측소로는 정확한 국지기상을 예측하는 데 한계가 있다. 대안으로 서울에서 운행 중인 택시를 활용해 보자는 것이다. 택시에 탑재한 ‘운행기록 자기진단장치(On Board Diagnostics, OBD)’의 센서를 통해 기온과 기압, 강수 등 외부 기상 정보를 실시간으로 획득할 수 있다.

정확한 관측을 통해 다량의 자료가 확보되면 이를 이용하여 예보를 생산하게 된다. 현재는 슈퍼 컴퓨터를 활용하여 생산된 수치 예보 자료가 기상 예보의 원재료다. 이제는 AI를 활용한 날씨 예보로 가야만 한다. AI는 오랜기간 동안 축적된 기압배치와 날씨 현황의 빅데이터 속에서 오늘과 유사한 기압계를 찾아낸다. AI 분석 자료, 수치 예보 자료와 인간 예보관이 상호 보완하는 예보는 획기적인 예보정확도 향상을 가져올 것이다. CNN(convolutional neural networks)은 빅데이터와 AI를 결합해 기상 영상 분석 능력을 15% 정도 향상시킨 것으로 알려져 있다. 이는 정말 엄청난 수준이라고 할 수 있다. 이러한 기술을 우리나라 기상 예보에 적용한다면 예보의 정확도를 향상시킬 수 있다. AI를 활용한 예측은 단기 예보 뿐만 아니라 산업계에 많은 도움을 주는 중장기 예보의 정확도도 향상시킬 것이다.

AI를 기상 예측에 결합한 일본과 IBM의 선진 사례

정확한 기상 정보의 첫 단계는 위험 기상에 대한 영향력 분석이다. 위험 기상(폭염, 한파, 태풍, 집중 호우, 폭설 등)이 사람과 산업에 미치는 최악의 시나리오를 AI와 빅데이터를 이용해 분석할 수 있다. 산출된 분석 정보들을 정확하고 빠르게 전달하는 데에는 IoT와 표시

기술을 활용할 수 있다. 지금은 지진, 산사태, 홍수 등 자연 재난 때마다 전화 문이나 인터넷 검색, 기상 안내 시스템이 마비되고 있다. AI를 활용해 이를 개선해 나간다면 통보 뿐 아니라 상담 서비스까지 가능해질 것이다. 일본 기상청의 지진 예측 및 감지, 전파 시스템이 그 좋은 예다. 이들은 빅데이터와 AI, IoT, 그리고 최고의 통신 기술을 결합해 지진 피해를 최소화하고 있다.

장기적인 기후변화 대응에 4차 산업혁명 기술은 다양하게 활용될 수 있다. IoT 및 빅데이터 기술을 활용하여 획득한 기후재난 데이터 및 지리·기후 정보 등의 상호연계분석을 수행한다. 이를 통해 특정 지역에서의 취약한 재난 유형을 미리 예측하여 조기 재난 경보에 활용할 수 있다. 또한 일사량, 강수량, 농작물 경작 현황 등을 종합적으로 분석하여 곡물 생산 계획에 반영한다. 이는 기후의 영향을 받는 임업, 수산업, 해양 등의 다양한 분야에 활용될 수 있다.

이처럼 4차 산업혁명 기술과 날씨를 잘 결합시킨 기업이 바로 IBM이다. IBM은 AI, 드론, 클라우드 플랫폼을 통해 기상 정보를 수집하고 예보 기반을 구축하고 있다. 또한 최단시간에 최상의 정보를 제공해주는 기상 정보 서비스 시스템이 구축되어 있다.

세계적인 전자 가전 회사인 일본의 파나소닉(Panasonic)도 기상 정보 서비스에 4차 산업혁명 기술을 적극적으로 활용하고 있다. "Tropical 4D를 활용한 파나소닉의 Global 4D 일기 예보 제품군을 통해 항공, 해운, 해상, 재생가능 및 탐사 에너지 시장, 보험 및 필수품 등 정부 및 날씨에 민감한 산업 내에서 파트너의 작업을 지속적으로 지원할 것입니다." 파나소닉 날씨 솔루션(Panasonic Weather Solutions) 관계자의 말이다. 파나소닉은 전 세계의 항공, 해운, 탐사에너지 시장 등에서 날씨 예보가 부정확하여 매년 수십 억 달러의 비용이 사라지고 있다고 분석했다. 이들은 빅데이터, 사물인터넷, 드론을 포함한 로봇기술, 슈퍼 컴퓨터를 활용한 모델과 인공지능을 활용하여 실시간으로 기상 정보를 제공하는 파나소닉 날씨 솔루션을 만들었다. 이 솔루션이 보유한 기상 모델링 플랫폼은 세계에서 유일하게 분야별 맞춤형으로 개발되었다. 기상 예측 기능은 4차원(경도, 위도, 고도 및 시간)의 상세한 대류권 데이터를 연속적으로 공급하는 등 파나소닉의 독점 대기 데이터 세트를 최대한 활용하고 있다. 파나소닉은 4차 산업혁명 기술을 활용하여 기상정보서비스를 획기적으로 발전시킨 대표적 사례다.

실외 미세먼지와 AI : 대만의 사례

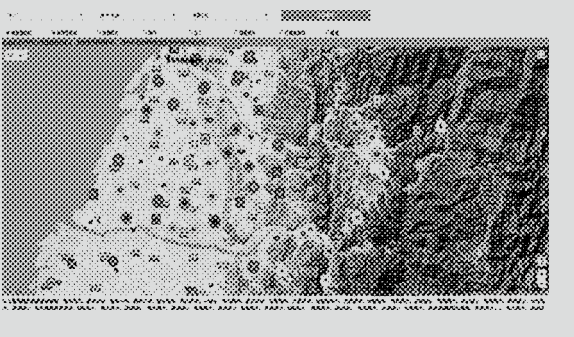
기후변화와 날씨는 미세먼지의 증가와 깊은 연관이 있다. "정작 미세먼지 비상대책이 발표된 날은 미세먼지 농도가 좋았습니다." 2017년 1월 14일 서울시는 미세먼지 비상대책을 발표했다. 바로 차량 2부제와 대중교통 수단인 버스와 지하철의 무료 탑승이다.

그런데 미세먼지 농도가 나쁠 것으로 예상해 대중교통을 무료로 제공했지만 정작 1월 15일 오전의 서울시 미세먼지 농도는 '보통' 단계였다. 이는 아직도 우리 나라 미세먼지 예측 기술이 낮다는 사실을 시사한다.

미세먼지에 대한 정확한 처방과 대책이 나오기 위해서 가장 시급한 것은 정확한 미세먼지 예측 능력이다. 이를 위해서는 관측 정확도의 향상과 관측소 증가가 필수적이다. 그런데 현행 환경부의 미세먼지 관측망의 숫자는 매우 적다. 초미세먼지 측정소는 2016년 4월에 152개소에 불과했다. 정부는 측정소의 개수를 2018년에는 287개소, 2020년에는 293개소로 늘리겠다고 한다. 측정 장비의 대폭 확장이 어려운 것은 장비가 고가(高價)이기 때문이다. 측정소의 개수가 절대적으로 부족한 것도 문제지만 측정소 가운데 높이 기준(1.5~10m)을 충족한 것은 전체의 26.9% 뿐이다. 높이를 충족하지 않은 관측소의 관측 값이 포함되었기 때문에, 정확한 미세먼지 관측 값을 얻기 어렵다. 여기에 공원(서울 성동구, 송파구)이나 정수시설(서울 광진구, 김포시 고촌면)처럼 지역 대표성이 없는 곳에 설치된 측정소도 많다. 그렇다보니 관측 값이 국민들이 체감하는 미세먼지 농도 값과 달라 불신감을 키우게 된다.

미세먼지는 한 지역의 좁은 곳에서도 농도 차이가 크다는 특성이 있다. 길 옆인지, 지하철 입구인지, 공장 옆인지, 아니면 공원인지 말이다. 따라서 현재 국가의 미세먼지 관측망으로는 내가 사는 지역의 미세먼지 정보를 정확하게 알 수 없다. 이런 문제를 해결하기 위한 방법에는 무엇이 있을까? IoT를 이용한 관측 자료 확장 방법이 있다. 간이 미세먼지 실외 측정기를 도시 곳곳에 설치하는 것이다. 미국 UC버클리 대학의 커크 스미스(Kirk Smith) 박사는 미세먼지 간이 측정기 이용에 적극적으로 찬성한다. "간이 측정기의 데이터를 통계에 활용하거나 정교한 수치를 필요로 하는 곳에서 사용하기는 어렵습니다. 그러나 저렴한 미세먼지 간이 측정기는 농도의 흐름을 정확하게 잡아낼 수 있습니다." 스미스 박사는 실외 미세먼지 간이 측정기들이 국가의 미세먼지 정책에 변화를 가져올 수 있다고 말한다.

【그림 2】대만의 실외 미세먼지 간이 측정기 설치 및 데이터 표출 그림

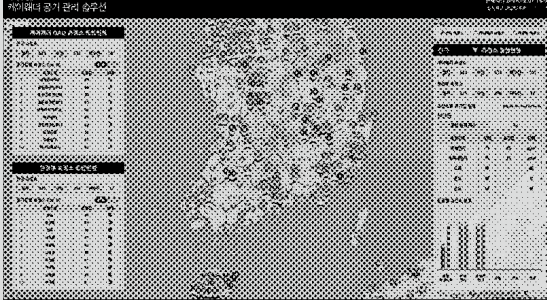


이런 방법을 도시에 활용한 나라가 대만이다. 대만의 기안 국제대학이 외부용 미세먼지 간이 측정기를 곳곳에 설치하였다. 이들은 관측 데이터를 대만 미세먼지 사이트를 통해 정부와 국민들에게 실시간으로 제공한다. 이를 통해 정부나 지자체는 각각의 다른 영역에 대해 초미세먼지의 저감 대책을 수립하고 실천하는 것이다. 이 덕분에 국민들은 최상의 미세먼지 정보를 제공받을 뿐만 아니라 정부가 미세먼지 저감을 위해 어떤 대책을 시행하는지에 대해서도 알 수 있다. IoT나 통신망을 활용한 대만의 수준은 아직은 초기 단계다. 그러나 대만은 데이터가 쌓이면 AI를 활용한 예측 단계로 나갈 계획이라고 한다.

미세먼지 측정에 AI가 적용된 국내 사례

국내에서도 케이웨더에서 실외용 미세먼지 간이 측정기를 만들어 현재 전국적인 관측망을 구성해 나가고 있다. 케이웨더는 대만의 사례와 유사한 간이측정기 관측망을 제주시에 구축했다. 62대의 실외 미세먼지 간이측정기를 설치하여 IoT 기반의 공기질 모니터링 서비스를 제공했다. 이 서비스를 통해 상세 공기질 측정 값을 제공하여 시민 건강 증진을 추구할 수 있게 되었다. 또한 실내공기 측정 값과 연계한 입체적 공기질 관리가 가능해졌고 지자체는 미세먼지에 즉시 대응이 가능해졌다. 케이웨더에서는 관측 자료의 데이터베이스 구축을 통해 대기오염과 건강의 상관성을 정량적으로 분석하여 빅데이터와 인공지능을 결합한 예보 체계로 나갈 계획이다.

【그림 3】케이웨더의 실외공기 측정소 분포 현황을 확인할 수 있는 모니터링 시스템

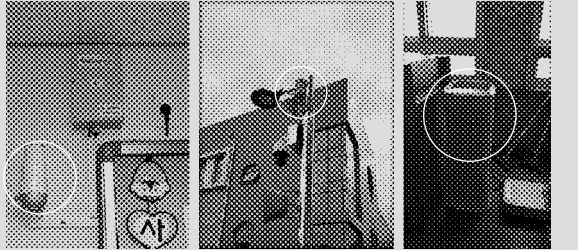


교육 현장에서도 미세먼지를 스마트하게 파악, 예측하는 시도를 벌이고 있다. 기실, 악화되고 있는 미세먼지에 피해를 많이 입는 곳이 학교다. 학교에는 고농도 미세먼지 대응 매뉴얼이 보급되어 있지만 적극적인 대응 조치가 미흡하다. 따라서 미세먼지 대응 교육을 체계적, 전문적으로 수행하는 선도학교 운영이 필요하다. 부산시 교육청에서는 IoT 기반 스마트 공기질 관리 체계를 구축하여 운영하고 있다. 부산시 교육청은 미세먼지의 유해성에 대한 경각심을 고취하고 학교 현장에 적합한 미세먼지 대응

매뉴얼을 마련함으로써 미세먼지 피해를 최소화하려고 한다. 부산시 교육청은 초등학교 5곳, 중학교 3곳, 고등학교 2곳 등 총 10개교를 미세먼지 대응교육 선도학교로 지정하여 운영하고 있다. 실내·외 공기 간이측정기를 학교 운동장, 교실, 체육관 등 실내/외 각 1대씩 설치하여 운영한다. 측정된 실내외 데이터를 종합적으로 분석하여 학습 능률 지수를 제공한다. 실외 미세먼지 간이 측정기를 활용하여 학교 체육이나 야외 활동 가능 여부와 야외 활동 지속 시간을 알려준다.

실내 미세먼지 농도가 높아지면 자동으로 공기청정기가 작동하도록 하였다. 이들 학교에서는 교육과정과 연계하여 학생과 교사를 대상으로 한 미세먼지 대응 교육을 추진 중이다. 미세먼지 국가관측망과 민간기업의 실외 공기 간이 측정기의 데이터를 연계하여 빅데이터화 하면 더 정확한 예측이 가능하다. 빅데이터와 AI가 결합되면 환기장치나 공기청정기가 자동으로 작동하고 학생은 물론 교사에게도 정보서비스가 실시간으로 전달될 것이다. 학생들이 최적의 공기 속에서 공부하고 생활하게 될 날이 멀지 않았다는 말이다.

【그림 4】부산 거성중학교 설치 운영 사례



교실에 실내공기측정기 (IAQ) 설치 운동장에 실외공기측정기 (OAQ) 설치 교실에 IoT 공기청정기 설치

*IAQ : Indoor Air Quality *OAQ : Outdoor Air Quality

지하철 내 미세먼지와 그리고 SI

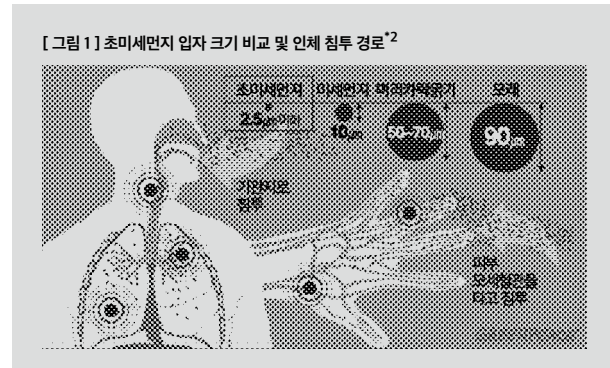
미세먼지 문제를 똑똑한 기술로 해결하기 위해서는 먼저 미세먼지가 무엇인지 정확히 이해할 필요가 있다. 최근 언론 보도를 통해 자주 언급되는 용어는 미세먼지 주의보, 황사 발생, 초미세먼지 발생, 2차 생성먼지 등 매우 다양하다. 이와 더불어 대기오염 외에 우리가 생활하는 공간의 실내 공기질에 대한 언급이 자주 등장하고 있다. 특히 하루 800만 명 이상이 이용하는 대중교통 시설인 지하철과 전국 590만 초중고생이 다니고 있는 학교, 유치원, 어린이집 등 생활 밀착형 공간에서의 미세먼지 문제가 중요한 화두가 되고 있다.

미세먼지란 무엇인가?

크기가 $10\mu\text{m}$ (마이크로미터, 1/1,000,000m)보다 작은 경우에 한하여 미세먼지(particulate matter less than $10\mu\text{m}$, PM10)라 칭하고 있다. 우리가 흔히 말하는 미세먼지는 공기 중에 부유하는 직경 $10\mu\text{m}$ 이하의 모든 액체 또는 고체인 물질(또는 티끌)이라 할 수 있다. PM2.5란 직경 $2.5\mu\text{m}$ 이하의 모든 공기 중 부유 입자를 말하는 것이다. 우리는 이를 미세먼지와 구분하여 초미세먼지라고 칭한다. $10\mu\text{m}$ 보다 큰 먼지는 입안의 점액, 콧털, 상기도에서 충분히 걸러지지만, $10\mu\text{m}$ 이하의 입자는 우리의 호흡기에 깊숙이 침투하여 폐까지 도달할 수 있다. 미세먼지가 기관지에 쌓이면 가래가 생기고 기침이 잦아지며 기관지 점막이 건조해지면서 세균이 쉽게 침투한다. 만성 폐질환이 있는 사람은 감염성 질환에 더욱 취약해지기도 한다. 특히, 초미세먼지는 폐포를 통해 혈관에 침투해 염증을 일으킬 수 있다. 이 과정에서 혈관이 손상되면서 협심증, 뇌졸중 등 심혈관 질환이 발생할 수도 있는 것이다.

최근 발표된 연구 결과에 따르면 초미세먼지로 인한 사망인구는 1999년 350만 명에서 2015년 420만 명으로 증가한 것으로 추정된다.¹ 미세먼지의 직경에 대한 연구와 인체 유해성 연구가 활발해지면서 학계에서는 폐의 더 깊은 부분(허파파리)까지 먼지가 침투하여 혈액과 산소 교환 시 인체에 해를 가하는 결정적 크기를 $2.5\mu\text{m}$ 로 한정했다. 이때부터 초미세먼지가 위험을 일으키는 먼지로 등극하게 됐다.

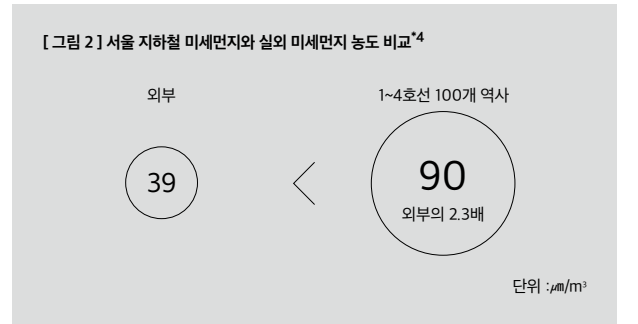
미세먼지의 크기 보다는 성분이 더 중요하지 않을까라는 의문에 대해서는, 크기가 조금 더 중요하다고 이해하면 좋을 것 같다. 앞서 언급했듯이, 크기에 따라 인체에 침투 여부가 결정되기 때문이다. 최근 몇 년 동안 대기 미세먼지 고농도 현상이 빈번하게 발생하면서 미세먼지에 대한 다양한 대책이 쏟아지고 있다. 특히 현대인이 대부분의 시간을 보내는 실내 공간의 공기질 관리 중요성이 증대되고 있으며, 이용자 수가 많은 지하 역사 등의 다중 이용 시설과 민감 계층이 이용하는 공공시설 및 학교 등의 미세먼지 저감이 시급하다.



지하철 공간의 미세먼지

저자는 실질적으로 국민들의 삶의 질을 높이기 위해 많은 사람들이 이용하는 지하철 미세먼지 문제를 인공지능을 통해 해결하는 방법에 대해 얘기를 풀어보고자 한다. 통상 지하철은 차량, 지하철역, 터널의 세 공간적 요소를 포함한다. 이 가운데 우리가 접하는 공간은 차량 내부나 지하철역 내부 공간인 대합실, 승강장, 환승 통로로 국한된다. 지하 공간은 기본적으로 부족한 일사와 높은 습도로 인해 공기질이 지상보다 전반적으로 좋지 않으며, 지속적인 환기(ventilation)를 통해 지상과 유사한 수준을 유지하는 것을 기본으로 하고 있다.

그러나, 환기만으로는 지하철 실내 공기질 문제를 해결할 수 없다. 왜냐하면, 지하철 공기질이 외기와 밀접하게 연관되어 있어 오염된 외기의 유입이 실내공기를 더욱 악화시킬 수 있기 때문이다.³ 특히, 지하철은 교통이 복잡한 도심의 지하공간에서 운행되기 때문에 도로에서 발생하는 매연 등 자동차 배출 오염 물질에서 자유로울 수 없다. 환기구와 출입구를 통해서 언제든지 도로 오염물질이 지하로 유입될 수 있다. 물론 외부에서 유입되는 공기 중 오염물질을 제거할 수 있는 저감장치가 일부 설치되어 운영되고 있지만, 고성능 사양이 적용되기는 어려운 실정이다.



지하철에서의 미세먼지 발생원은 차량 운행으로 발생하는 마모 입자, 열차풍에 의한 재비산(再飛散)과 외부에서 유입되는 미세먼지라고 볼 수 있다. 차량 운행으로 발생하는 마모 현상은 차륜(wheel)과 선로(rail) 간 마찰, 전력선과 집전부(pantograph) 간 마찰, 브레이크 마찰로 구분되며 이 모든 마찰은 다양한 크기의 먼지 또는 초미세먼지의 발생원이 되고 있다.⁵

지하철 미세먼지에 대해 기본적인 이해가 되었다면, 지하철 미세먼지 문제 해결에 SI를 활용하는 시도를 해보자. 우선 지금까지의 내용을 정리해보자. 첫째, 지하철 먼지는 외부에서 유입되거나 터널 안에서 차량 운행에 따라 발생하고 있다. 둘째, 미세먼지를 포함한 나쁜 공기를 정화하려면 환기가 필요하지만 환기량이 증가하면 운영 비용이 증가한다. 셋째, 환기량을 증가시켜도 바깥 공기가 좋지 않을 경우 지하철 공기질은 개선되지 않는다. 이 세 가지 문제를 해결할 수 있는 간단한 방법은 아래와 같다.

글 | 권순박 dapairsolution@gmail.com

2003년 미세먼지 분야 박사 학위 취득 후 서울대학교와 일본기계연구원(AIST)을 거쳐, 2006년부터 한국철도기술연구원 교통환경연구팀에서 지하철 미세먼지 문제를 연구하고 있습니다. 지난 2017년 9월 '인공지능 기반 미세먼지 해결'이라는 비전을 가지고 벤처기업 '딥디에이피(DAP: Deep-learning Air solution for Public)'를 창업하였습니다. 다양한 연구 개발 과제를 수행하면서 논문과 특허 위주의 연구를 수행하던 정부출연연구기관 연구원이 스타트업 벤처의 대표가 된 이후, 기술실용화를 통한 사업화라는 새롭고 어려운 도전을 이어나가고 있습니다. 이 자리를 빌어 쉽지 않은 도전을 함께 하고 있는 저희 디에이피 직원들에게 감사를 표하고 싶습니다.

'바깥 공기가 좋을 때만 환기를 실시하여 비용도 절감하고 공기질도 개선한다.'

그러나 이 해결 방안에는 외부 공기상태가 계속 좋지 않으면 환기를 할 수 없는 상황에 처하게 된다는 약점이 있다. 지하철 안에는 계속 먼지가 발생되고 있는데 말이다. 따라서, 환기를 하지 않고도 공기질을 개선할 수 있는 공기정화장치가 필요하다.

이제 앞선 간단한 답에 조건을 추가하여 다음과 같은 방법을 적용할 수 있다.

'바깥 공기가 좋을 때는 환기를 하고, 좋지 않을 때는 실내 공기정화장치를 가동하는 것이다.'

여기에 기계 설비(환기 장치, 정화 장치 등) 운영의 문제를 덧붙이자면, 공기를 처리하여 효과가 나타나기까지는 일정 시간이 소요된다. 즉, 공기에 먼지가 많을 때는 기계 설비를 가동시켜도 효과가 나타나기까지 시간이 제법 소요되고 처리 공간이 커질수록 더 많은 시간이 걸린다는 것이다. 따라서 기계 설비의 성능이 제대로 구현되려면 사전에 설비를 가동시킬 필요가 있다. 미세먼지 오염 상태를 미리 파악할 수 있다면 말이다. 위의 사항을 반영하여 지하철 미세먼지 문제를 아래와 같은 방법으로 해결할 수 있다.

'바깥 공기 상태와 실내 공기 상태를 이용하여 지하철 미세먼지 오염을 예측하고, 이에 맞추어 환기나 공기정화 설비를 사전에 가동한다.'

지하철 미세먼지 문제를 해결하기 위한 AI

지하철 미세먼지 문제를 해결할 수 있는 방법의 핵심은 여러 데이터를 분석하여 미세먼지를 사전에 예측하는 것이다. 이때 AI를 활용하면 가장 정확한 예측이 가능할 것이다. 그렇다면 어떤 데이터를 통해 지하철 미세먼지를 예측할 수 있을까? 가장 좋은 방법은 바깥 공기 상태와 실내 공기 상태를 직접 측정하여 데이터를 얻는 것이다. 그러나 정확한 측정을 위해 많은 비용과 노력이 요구되기 때문에 보다 효율적인 데이터 수집 방안을 고민할 필요가 있다. 정부에서 제공하는 다양한 공공 데이터는 가장 효율적인 데이터 중 하나이다. 에어코리아(AirKorea)는 전국 97개 시군에 설치된 323개의 측정망에서 1시간 단위로 미세먼지를 포함한 다양한 오염 물질의 농도 값을 측정하여 공공 데이터를 제공한다.

미세먼지 농도에 영향을 주는 또 다른 요소는 기상 조건이다. 기상정보에 대한 공공데이터는 기상청에서 제공하고 있으며 그 항목에는 기압, 기온, 풍향, 풍속, 습도, 강수량 등이 있다. 이 외에 추가로 활용할 수 있는 데이터는 도로 교통량 정보다. 도심 지역에 위치한 지하 역사는 환기구 주변 도로 교통 오염원에서 자유로울

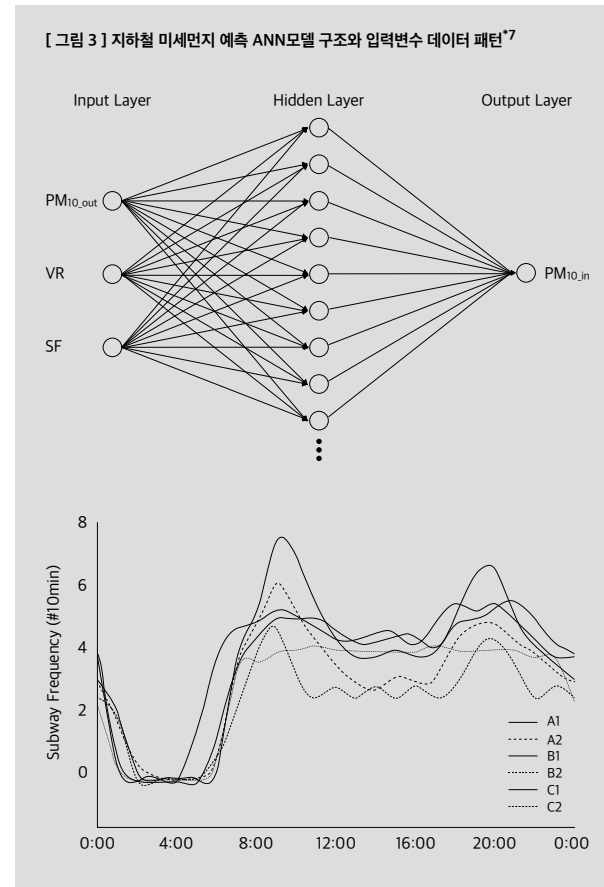
수 없기 때문에 교통량이 많은 곳 또는 정체가 심한 곳에서 매연 등의 오염 물질이 많이 발생할 것으로 추측할 수 있다. 따라서 해당 역사 인근 교통량 정보는 중요한 데이터가 될 수 있다. 외부 공기의 정보를 유추할 수 있는 공공 데이터는 이렇게 세 가지 요소 정도로 요약할 수 있다.

그렇다면 지하 역사 안에서의 실내 공기질 상태는 어떻게 파악할 수 있을까? 앞서 언급했듯이 지하철에서 미세먼지를 발생시키는 원인은 차량의 운행이다. 따라서 차량의 운행 횟수와 실내 공기질과는 상관관계가 존재할 것이다. 한편, 기계설비 운영 상태 데이터를 활용하여 실외공기와 실내공기가 교환되는 환기 정도를 파악할 수 있다. 우리는 오랜 기간 지하철 역사에서 실내 공기질을 측정해오고 있으며 지금도 미세먼지 농도와 실내오염도 등을 꾸준히 모니터링하고 있다. 이러한 데이터를 기계학습을 위한 참값 데이터로 활용할 수 있다.

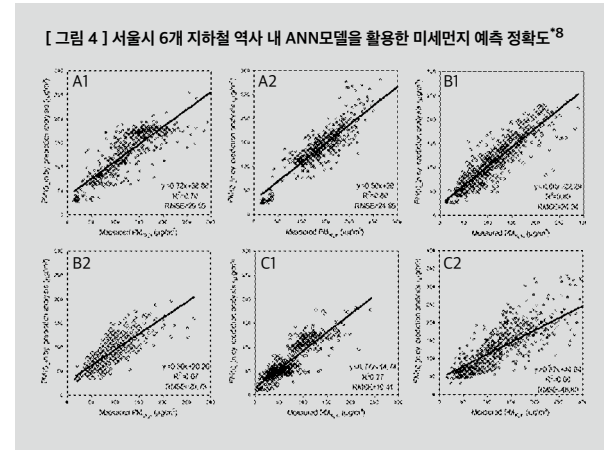
Features : 대기오염 정보, 기상 정보, 교통량 정보, 차량 운행 횟수, 공조 환기 설비 운영 정보 등
Labels : 지하 역사 미세먼지 농도, 지하 역사 초미세먼지 농도

미세먼지 또는 초미세먼지의 농도를 상기 특징(features)을 통해 학습시키는 인공지능 알고리즘은 영상처리나 자연어처리 등의 알고리즘에 비하여 비교적 단순하여, 복잡한 연산과정 없이 단층 신경망(single-layer neural network) 또는 층(layer) 몇 개를 추가한 다층 신경망(multi-layer neural network)으로 약 70% 이상의 정확도를 달성할 수 있다. 실제로 저자가 ANN(artificial neural network)을 이용해 외부 공기 환경 데이터로 지하철 실내 미세먼지를 예측하는 모델을 구축해 보았고, 이 모델이 67~80%의 정확도를 나타 낸다는 연구 결과를 발표한 바 있다.*6

지하철 미세먼지 예측을 위한 가장 심플한 ANN 모델은 외기 미세먼지(PM10_out), 환기량(ventilation rate, VR), 지하철 차량 운행 스케줄(subway frequency, SF)의 세 가지 변수를 통해 지하철 역사 승강장에서의 미세먼지 농도(PM10_in)를 예측하는 모델이다. 다음 그림은 ANN 모델 구조와 시간당 지하철 차량 운행 스케줄 데이터 패턴이다.



모델에 적용된 학습 알고리즘(learning algorithm)은 모멘텀 역전파(momentum backpropagation)가 적용된 경사하강법(gradient descent) 방식이며, 패턴 검색(pattern search)을 통해 학습 속도(learning rate, Lr) 및 노드(node)의 최적 개수를 도출하였다. 전체 데이터의 80%를 학습(training)에 사용하였고, 나머지 20%의 데이터를 검증(validation)에 사용하는 교차 검증(cross validation) 방식을 적용하였다.



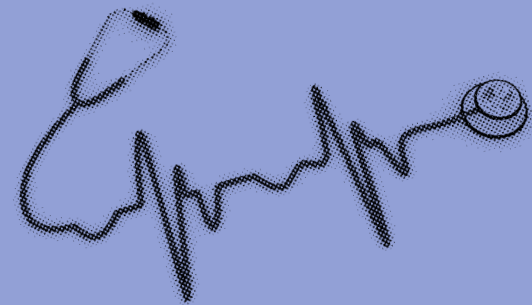
6개 지하 역사(A1, A2, B1, B2, C1, C2)에 대해 역사별로 분석을 수행하였으며, 실제 측정 값과 예측 값의 상관 관계를 [그림 4]에

제시하였다. 예측 값의 정확성은 각기 다른 특성을 가진 6개 역사 별로 다소 차이가 발생하는데, 예측 값과 실측 값간 상관성(R²)은 약 0.67~0.80 수준으로 나타났으며 해당 역사의 심도(깊이)가 깊어 질수록 상관성이 낮아지는 것으로 분석되었다.

앞서 이야기한 지하철 미세먼지 농도 예측 결과는 간단한 입력 변수만을 통해서도 승강장 미세먼지 농도를 예측할 수 있다는 것을 보여준다. 다양한 입력 변수를 사용하거나 개선된 알고리즘을 적용한다면 실제 미세먼지 예측 결과의 정확도를 향상시킬 수 있을 것이다. 또한 실제 지하철 미세먼지 문제를 해결하려면 단순한 공기질 개선 문제 외에도 운영 비용 최소화, 이용객 온열 쾌적성 향상, 여름철 피크타임 전력컨트롤 등 다양한 문제에 대한 해답이 필요하다. 이를 위해서는 외부 공기 데이터, 실내 환경 데이터, 설비 운영 데이터 등 현재 지하 역사의 다양한 정보를 활용하여 향후 미세먼지 농도를 예측할 수 있는 인공지능 기술의 개발이 필요하다. 또한 미세먼지가 기준 값을 초과할 것으로 예상되면 미세먼지 농도를 저감할 수 있는 기계설비의 운영방식을 제시하고 저감된 농도를 다시 예측하는 AI기술 적용이 필요하다. 이러한 두 단계의 인공지능 분석을 통해 지하 역사의 미세먼지 농도를 예측하고 필요에 따라 스스로 판단하여 역사 공조설비 및 정화장치를 가동시키는 기술은 현재 지하철 역사의 단순한 스케줄 제어로는 대응할 수 없는 혁신적인 스마트 제어를 가능하게 할 것이다. 이를 통해 지하철 미세먼지에 대한 효율적 관리는 물론 설비 운영 에너지도 절감할 수 있을 것이다. 향후 AI 기반의 지하 역사 공조설비 운영기술은 미세먼지 문제 대응 핵심기술로 활용될 것이며, 더 나아가 다양한 공공 시설물에도 확산될 것이 자명하다.

*1 논문 | Cohen, J., et al. (2017). Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015, The Lancet 389.10082, 1907-1918 *2 참고 | 그린피스 <http://slownews.kr/38558> *3 논문 | Kwon, S.B., et al. (2015). A multivariate study for characterizing particulate matter (PM10, PM2.5, and PM1) in Seoul metropolitan subway stations, Korea, Journal of hazardous materials 297: 295-303. *4 참고 | 연합뉴스 <http://www.yonhapnews.co.kr/digital/2017/07/05/4905000000AKR20170705075300797.HTML> *5 논문 | Namgung, H.G., et al. (2016). Generation of nanoparticles from friction between railway brake disks and pads, Environmental science & technology 50.7: 3453-3461. *6 논문 | Park, S., et al. (2018). Predicting PM10 concentration in Seoul metropolitan subway stations using artificial neural network (ANN), Journal of hazardous materials 341: 75-82. *7 논문 | Park, S., et al. (2018). Predicting PM10 concentration in Seoul metropolitan subway stations using artificial neural network (ANN), Journal of hazardous materials 341: 75-82. *8 논문 | Park, S., et al. (2018). Predicting PM10 concentration in Seoul metropolitan subway stations using artificial neural network (ANN), Journal of hazardous materials 341: 75-82.

의료 AI의 역사 그리고 발전 과제 part.1



AI & 의료 섹션에는 의료 진단을 소재로 AI가 의료에 적용된 역사에 대해 훑는 글을 첫 번째로 섭외했습니다. 이 글에는 병원에서 자주 접하는 ‘자동 심전도 판독’ 기술의 발전 역사와 관련된 AI 이야기를 담았습니다. 두 번째 글에는 AI를 포함한 정보통신기술 기반의 ‘헬스케어’ 시대 구현을 위해 확립해야 할 사회적 기반과 가치를 고민한 내용을 담았습니다.

의료 AI의 안착 과정

자동 심전도 판독(Automated ECG interpretation)의 발전 역사를 통해

글 | 정세영 syjung@snuh.org

의료 인공지능에 대한 대중들의 관심이 높아지고 있습니다. 의료 자원의 낭비를 막고 치료의 질을 높일 수 있는 방법에 대해 인공지능기술을 통해 찾는 연구자도 급속하게 늘어나고 있습니다. 그렇다면 그동안 의료 분야에서는 인공지능기술이 전혀 쓰이지 않았을까요? 역사적인 실례를 통해 인공지능 의료기기 같은 신의료기술이 태동하고 시장에 안착되기까지 어떤 과정을 거치는지 살펴봄으로써 의료 인공지능 연구자들이 놓치지 말아야 할 교훈에 대해 이야기 나누고 싶습니다.

응급실 스냅샷

2007년 11월 21일 오전 4시. 서울. 제법 매서워진 찬바람은 겨울이 오고 있음을 알린다. OO대학교 병원 응급실 인턴 000 선생은 초점이 나간 눈으로 컴퓨터 화면을 응시하고 있다.

‘이제 이 환자만 보면 조금 쉬 수 있겠구나’

태풍이 휘몰아치듯 환자들이 몰려 들다가 새벽이 되자 잠잠해졌다. 대기 환자 명단에 남은 한 명의 환자는 평온해진 응급실을 대변하고 있다.

67세, 여성.

“저녁 식사를 한 후 계속 속이 답답하네요”

별다른 과거 병력이 없던 환자는 몸이 불편한 듯 표정이 좋지 않았다. 그리고 본인의 증상을 조금씩 이야기하기 시작하였다.

‘CC> Abdominal discomfort, onset: 8 hours ago, 금일 저녁 식사를 한 후 prev. healthy...’

환자 보호자는 대수롭지 않은 증상으로 응급실을 따라온 것이 못마땅한 듯 보였다. 000 선생은 쏟아지는 졸음을 참으며 자판을 두드렸다. 그리고 채근하던 보호자의 말을 애써 무시하면서 신체 진찰을 하였다. 급성 복증(acute abdomen)^{*1}에 합당한 징후는 없었다.

‘아.. 이제 검사 결과만 확인하면 끝이다! 좀 쉬자...’

혈액검사, 복부 엑스선 검사, 심전도 검사 오더를 내렸다. 얼마 후 검사 결과가 나왔다

‘검사 결과 별 이상 없겠지?... 잘 설명해서 보내야겠다...’

졸린 눈을 비비면서 검사 결과를 확인한 000 선생은 심전도 결과지를 보며 정신이 번쩍 들 수밖에 없었다.

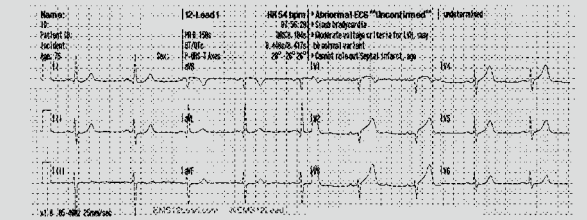
Alert! ACUTE MYOCARDIAL INFARCTION^{*2}

막 나온 혈액 검사 결과를 자세히 보니 심근 관련 효소 수치도 올라가 있었다. 그는 재빠르게 담당 백업 레지던트를 호출했고, 심혈관조영실 당직의에게도 전화를 했다. 그렇게 그 환자는 응급실을 떠나 심혈관조영실로 올라가게 되었다.

심장은 우심방 꼭대기에 있는 동결절에서 생성된 전기 자극에 의해 분당 60~100회의 속도로 규칙적으로 박동한다. 심전도(心電圖, electrocardiogram, ECG)는 이런 심장의 전기 활동을 측정하는 방법으로 피부에 부착된 전극과 신체 외부의 장비로 기록된다. 심장질환이 있으면 전도 조직이 손상될 수 있으며 이는 심전도의 변화로 나타날 수 있다.^{*3}

현재 병원에서 사용되는 심전도 기기는 판독 결과를 수초 내로 제공한다. 이를 ‘자동 심전도 판독(automated ECG interpretation)’이라고 한다.

[그림 1] 자동 심전도 판독 결과의 예시^{*4}



Normal ECG “Unconfirmed”

“Unconfirmed”는 최종 진단을 내리기 전의 상태란 의미로, 의사가 환자의 임상 정보를 통합하여 판독하는 ‘overread’라는 과정을 거쳐야 최종 결과가 도출된다.

의대생들은 심장의 전기생리학부터 심전도의 원리와 판독방법까지 배운다. 병원 실습을 시작하면 심전도가 의료 현장에서 실제로 어떻게 활용되는지 경험하게 된다. 의사로서 처음 발을 내딛는 인턴 초반에는 심전도 기기의 각 유도(lead)를 몸의 어느 부위에 어떻게 붙여야 노이즈 없이 정확한 검사를 얻을 수 있는지 수없이 반복 학습하게 되지만, 나중에는 자동 심전도 판독 결과를 각각의 임상 상황에서 어떤 의미로 해석해야 하는지 체득하게 된다.

예컨대 자동 심전도 판독 결과 정상심방리듬(normal sinus rhythm)이 나오면 대부분 상황에서는 안심할 수 있다. 하지만 흉통을 지속적으로 호소하는 환자의 경우 ST분절 비상승 심근경색증(Non-ST elevation myocardial infarction, NSTEMI)을 배제할 수 없기 때문에 심전도 결과 이상이 없더라도 안심할 수 없다. 심전도는 대부분 상황에서 음성예측도^{*5}가 높다.^{*6} 즉, 심전도 결과가 정상으로 나오면 대부분 실제로도 정상이고 안심해도 된다는 의미이다. 양성예측도^{*7}는 질환에 따라 달라진다. 예컨대 심전도에서 급성심근경색(acute myocardial infarction)이나 방실전도차단(atrio-ventricular conduction block)^{*8}에 대해서는 양성예측도가 낮으므로 심전도 판독에 해당 질환이 있더라도 실제 그 질환이 있는지 임상이는 꼼꼼하게 점검해야 한다.

몇몇 상황에서 한계가 있기는 하지만, 대체로 자동 심전도 판독은 매우 고마운 기술이다. 특히 초보 의사에게는 더욱 그렇다. 그리고 임상 현장에서 널리 사용된다. 심전도를 찍으면 자동판독은 ‘당연히’ 나오는 결과다. 또한, 의사로서 경험이 쌓일수록 그 결과가 어느 상황에서 어떤 의미를 갖는지 정확하게 알게 되고 적절히 활용하게 된다.

그렇다면 자동 심전도 판독은 어떤 기술일까? 위키피디어를 살펴 보도록 하자.

자동 심전도 판독은 인공지능과 패턴인식 소프트웨어, 지식 베이스를 활용한 기술로 환자의 심전도 결과를 분석하여 컴퓨터 기반 진단 및 해석을 자동으로 제공한다.^{*9}

즉, 자동 심전도 판독은 인공지능 기술이다. 조금 더 살펴보자.

자동 심전도 판독 과정*10

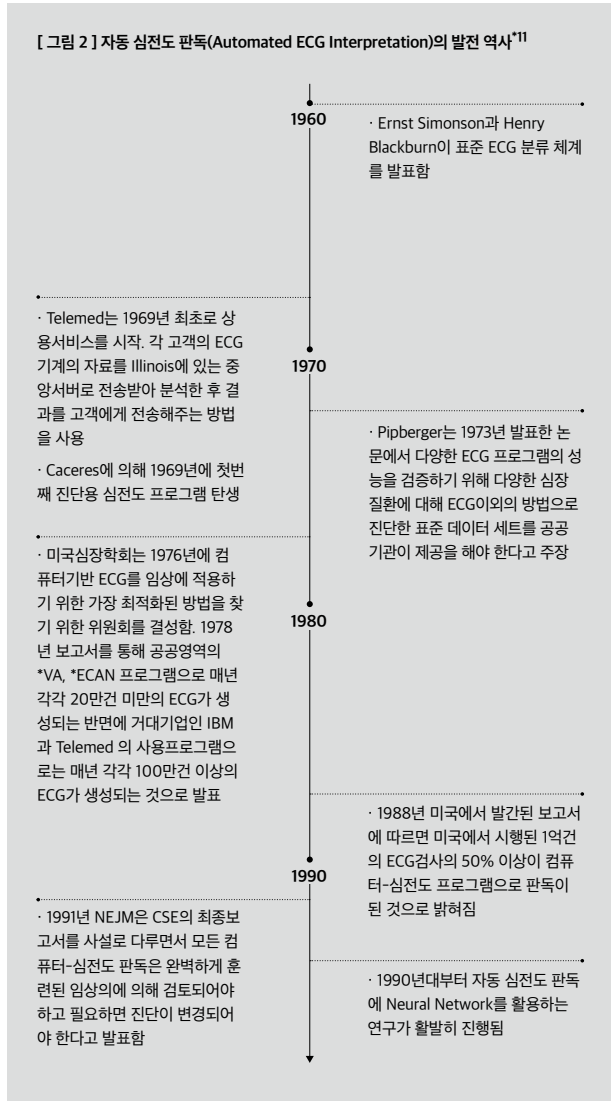
1. 각 ECG 채널에서 얻은 아날로그 신호를 전용 소프트웨어와 DSP(digital signal processing) 칩을 이용해 디지털 신호로 변환한다
2. 변환된 디지털 신호를 일련의 알고리즘을 이용해 처리한다. 예컨대 노이즈 제거, 기저치 변동 제거 등이다.
3. 특징 추출(feature extraction): 수학적 방법을 사용하여 해석 및 진단에 중요한 분석을 시행한다. 분석결과는 AI 기반 알고리즘에 입력된다. 패턴분석 프로그램에 정보가 입력되기 전에 푸리에(Fourier) 분석이나 파형 분석같은 과정이 선행된다.
4. 규칙기반 전문가 시스템(rule-based expert system), 베이지안 확률 분석(probabilistic Bayesian analysis), 퍼지 논리 알고리즘(fuzzy logic algorithms), 클러스터 분석(cluster analysis), 인공 신경망(artificial neural networks), 유전 알고리즘(genetic algorithm) 등을 사용해서 분석을 한다.
5. 결과를 보고한다.

심전도를 찍으면 위와 같은 과정이 수초 이내에 진행되어 바로 결과가 출력된다.

자동 심전도 판독은 언제부터 개발되었고 사용되기 시작하였을까? 의사들조차도 대부분 심전도 기기가 인공지능 기술인지 알지 못한다. 의료 현장에 이처럼 완벽히 이식(implementation)된 인공지능 의료기기가 있을까? 그리고 인공지능 기술인 자동 심전도 판독 기기의 발전 과정을 보면 인공지능 기반 신의료기기가 의료 현장에 정착되기까지 어떤 과정이 필요한지 힌트를 얻을 수 있지 않을까?

본 글에서는 먼저 자동 심전도 판독 기술의 발전 역사를 찬찬히 짚어 보면서 새로운 의료기기가 의료 현장에 정착되기까지 어떤 과정을 거치게 되는지 알아보겠다. 다음으로 왜 그런 과정을 거칠 수 밖에 없었는지에 대해 의학 연구와 임상 의사 결정의 특징을 고려하여 살펴보도록 하겠다.

자동 심전도 판독의 발전 역사



1) 1960년대 : 자동 심전도 판독 기술의 태동 시기

- 에른스트 시몬슨(Ernst Simonson)와 헨리 블랙번(Henry Blackburn)이 1960년 표준 심전도 분류 체계를 발표하였다.
- 허버트 핍버거(Hubert Pipberger)는 재향군인회 프로젝트를 통해 8개의 재향군인 병원에서 ECG-independent Frank data를 수집하여 VA 프로그램을 개발하였다. 미국 국립보건원(National Institutes of Health)에 있었던 세사르 카세레스(Cesar Caceres)는 표준 12-유도 심전도를 기반으로 하는 심전도 판독 프로그램인 ECAN을 개발하였다.
- 초기에는 ECAN 프로그램 도입률이 저조했는데, 단일 채널(single channel) 기록으로 나온 12-유도 심전도 결과를 해석하는 것이 효율적이지 못했기 때문이다.
- 핍버거(Pipberger)의 VA 프로그램도 사용률이 낮았다. 베이즈 방법 기반의 진단 분류 기준이 여러 진단을 조합하는 것을

허용하지 않고 상호 배타적인 진단만 허용했기 때문이다. 실제로 임상에서 심장전문의는 심전도 결과에 환자의 임상정보를 결합하여 진단의 정확도를 높이는 확률론적 방법을 사용한다.

2) 1970년대 이후 : 기술 발전 및 고도화 시기

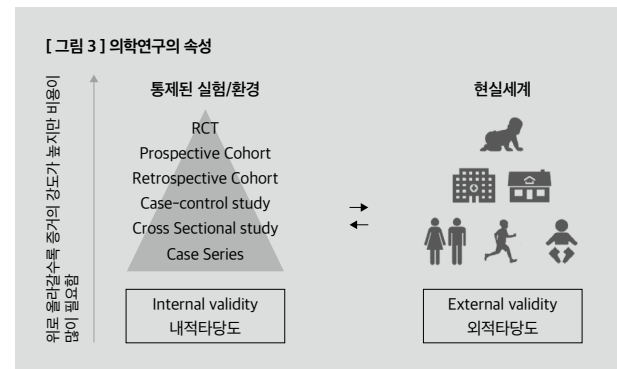
- 메이오 클리닉(Mayo clinic)의 랄프 스미스(Ralph Smith)와 IBM은 합작 벤처회사를 설립하여 표준 12-유도 ECG와 변형 Frank 유도 방법을 통합하는 연구를 진행하였다.
- Mayo-IBM 합작사의 프로그램은 IBM 360 컴퓨터에서 구동되었고 후에 my ECG research center에 의해 IBM 1800 컴퓨터에 이식되었다. 향후에는 IBM 이외의 컴퓨터에도 사용 가능하게 되었다.
- 레이 보너(Ray Bonner)는 다기관협력프로젝트를 통해 IBM 5880을 사용하여 카드 장착용 ECG프로그램을 개발하였는데 이는 심전도 해석 결과를 저장하고 출력하는 것이 최초로 가능해진 획기적인 장치였다. 그 장치에 설치된 컴퓨터 기반의 ECG 프로그램은 텔레메드(Telemed), 마켓(Marquette), HP 필립스(HP-Philips), 모르타라(Mortara)가 개발한 제품이었다.
- 레이 보너는 자동 심전도 판독이 실용적이고 정확할 수 있다는 점을 최초로 증명하였다.
- 여러 연구자들의 노력으로 자동 심전도 판독 프로그램은 매우 빠른 속도로 발전했지만, 임상에서의 활용은 더딘 속도로 진행되었다.
- 미국심장학회는 1976년에 자동 심전도 판독 기술을 임상에 가장 효율적으로 활용하기 위한 위원회를 결성하였다. 또한 1978년 보고서에서 공공영역에서는 VA(Veternas Affairs), ECAN(Esophageal Cancer Action Network) 프로그램으로 매년 각각 20만 건 미만의 ECG가 생성되는 반면에 거대 기업인 IBM과 텔레메드의 사용 프로그램으로는 매년 각각 100만 건 이상의 ECG가 생성된다는 조사 결과를 발표하였다.
- 텔레메드는 1969년 최초로 상용 서비스를 시작하였는데 각 고객의 심전도 기계의 자료를 일리노이에 있는 중앙 서버로 전송받아 분석한 후 결과를 고객에게 전송해주는 방법을 사용하였다.
- 1978년까지 텔레메드는 1600명 이상의 고객을 확보하였고 매달 12만 5천여 건 이상의 심전도를 분석하게 되었다
- 1960~70년대에는 수많은 심전도 판독 프로그램이 개발되었다. Pipberger는 1973년 발표한 논문에서 공공기관이 표준 데이터 성능을 평가할 때 쓰일 수 있는데, 심전도 이외의 방법으로 진단된 심장질환에 대한 정보가 포함되어야 한다.
- 데이비드 모르타라(David Mortara)는 1970~80년대에 컴퓨터 기반 자동 심전도 판독 기술의 발전에 지대한 영향을 미쳤다.

- 심전도 판독 기술을 검증하기 위한 표준 데이터셋이 만들어졌다. 유럽 위원회(European commission, EC)는 ‘정량적 심전도 분석을 위한 공통 표준(common standards for quantitative electrocardiography, CSE)’ 데이터셋을 만들었다. 조스 윌렘스(Jos Willems)는 CSE 프로젝트를 통해 9개의 ECG 판독 프로그램과 8명의 심장 전문의의 판독을 비교 분석하였고 심장 전문의 대표 8인의 판독은 심전도 프로그램의 평균보다 5.8% 더 정확한 것으로 드러났다. (75.5% vs 69.7%, p<0.001)
- 최종 CSE 보고서에는 컴퓨터 기반 심전도 프로그램이 임상 의사가 지속적이고 일관된 진단을 내리는데 도움이 된다는 내용이 담겼다.

3) 1980년대 이후 : 인공지능 기술 도입 시기

- 1988년 미국에서 발간된 보고서에 따르면 미국에서 시행된 1억 건의 심전도 검사 결과의 50% 이상이 자동 심전도 판독 프로그램에 의해 보고되었다.
- 1991년 뉴잉글랜드 의학저널(The New England Journal of Medicine)은 CSE의 최종 보고서를 종설로 다루면서 모든 컴퓨터 기반 심전도 판독은 완벽하게 훈련된 임상 의사에 의해 검토되어야 하고 필요하면 진단이 변경(overread)되어야 한다고 발표하였다.
- 컴퓨터 기반 심전도는 serial ECG comparison과 같은 분야의 발전을 가져왔다.
- 1990년대부터 자동 심전도 판독에 신경망(neural net)과 같은 인공지능 기술을 도입하는 연구가 활발히 진행되기 시작하였다.
 - 위에서 살펴본 바와 같이 자동 심전도 판독은 지난한 발전의 역사를 갖고 있다. 왜 그래야만 했을까? 의료 신기술이 임상에 널리 활용되기 위해서는 어떤 과정들을 거쳐야만 할까? 이를 정확하게 이해하기 위해서는 의학 연구와 임상 의사 결정 과정의 속성을 올바르게 이해할 필요가 있다.

의학연구의 속성



의학은 기본적으로 사람을 치료하기 위한 학문이다. 따라서 증거를 창출하여 임상에 적용하기까지의 과정이 신중할 수밖에 없다. 아스피린이라는 약물을 개발하여 효과가 있는지 검증하는 연구를 생각해보자. 개발된 약을 아무런 검증 없이 환자에게 바로 적용한다고 가정해보자.

열나는 환자에게 아스피린을 사용했다. 환자는 열이 떨어졌다. 그렇다면 그게 과연 아스피린의 효과라고 할 수 있을까? 우리는 경험적으로 열이 날 때 물을 충분히 마시고 옷을 벗어도 체온이 떨어질 수 있다는 점을 알고 있다. 이렇듯 아스피린을 먹은 환자가 열이 떨어진 것이 아스피린의 순수한 효과인지 알기 위해 아스피린을 사용한 경험적인 사실을 수집하는 것은 강력한 증거가 될 수 없다. 교란 요인¹²으로 인해 열이 떨어졌을 수도 있기 때문이다. 아스피린의 효과에 대한 정확한 증거를 얻기 위해서는 통제된 환경에서 임상시험을 해야 한다. 아스피린과 똑같이 생긴 가짜 약을 만들고 무작위로 나누는 실험군(아스피린을 먹는 그룹)과 대조군(가짜 약을 먹는 그룹)에 각각 아스피린과 가짜 약을 준 후 그 효과를 비교하는 것이다.

실험군과 대조군을 무작위로 나누면 각 그룹에 특정 성별, 나이를 갖는 환자가 몰리거나 특정한 임상적 특징을 공유하는 환자가 몰릴 가능성이 떨어지게 된다. 즉, 아스피린 복용 여부를 제외한 다른 특징은 두 군에서 동일하게 되기 때문에 위 실험의 결과는 순수하게 아스피린 복용 유무에 따른 차이라고 볼 수 있다. 또한 약을 먹는 사람이 실험군인지 대조군인지, 주는 약이 아스피린인지 가짜 약인지 연구자도 모르게 하면 실험이 더욱 완벽해진다. 이를 이중 맹검 무작위 대조군 연구(double blind randomized controlled study, double-blind RCT)라고 하며 의학에서 가장 강력한 연구 방법이다.

연구 결과로 아스피린의 해열 효과가 입증되었다고 하자. 연구자들과 의사들은 그 결과를 어떻게 신뢰할 수 있을까? 이를 위해서는 연구결과로 발표한 논문의 방법론을 검토해야 한다. ‘교란 요인을 제대로 통제했는가?’, ‘강력한 증거 능력을 갖는 연구 설계를 했는가?’, ‘피험자는 충분한가?’와 같은 질문을 하게 된다. 연구 방법론은 증거의 능력이 가장 약한 증례 보고부터 가장 강력한 이중 맹검 무작위 대조군 연구까지 다양하다. 이렇게 통제된 환경에서 약물이나 치료의 효과가 입증되면 ‘내적 타당도(internal validity)’가 입증되었다고 말할 수 있다.

이제 아스피린이라는 신약이 출시되었다. 제약회사는 ‘이중 맹검 무작위 대조군 연구’에서 효능을 입증 받은 획기적인 약물이라고 선전한다. 임상 의사들은 환자의 병마와 싸우는데 사용할 새로운 무기가 생겼다고 기뻐한다. 아스피린은 불티나는 속도로 팔리기 시작한다. 대부분의 환자에게서는 효과가 있었다. 하지만

사람마다 효과의 차이는 조금씩 달랐으며, 임상 시험에서는 보이지 않았던 부작용들이 조금씩 보고되기도 한다. 하지만 아스피린이라는 약을 지속적으로 사용하는 데는 문제 없는 것들이다. 아스피린을 생산하는 제약회사의 주가는 계속 오르게 된다.

이와 같이 현실의 약물 사용 경험은 반드시 필요하다. ‘외적 타당도(external validity)’를 입증하기 위해서다. 내적 타당도는 흡사 실험실 환경처럼 철저히 통제된 조건에서 나온 결과라서, 새로 개발된 약물은 다양한 특징을 갖는 환자들이 산재해 있는 현실에서도 효과가 있는지 확인해야 한다. 실제로 내적 타당도는 입증했지만, 출시 후 예기치 않는 부작용이 나타나서 퇴출당하는 약물들도 있다.

다른 예를 살펴보자. 이번에는 현실 세계 사용 경험으로 의심을 품고 실험 연구로 증명하는 예다. 폐경 후 호르몬 대체 요법(postmenopausal hormonal replacement therapy, HRT)은 부족한 여성 호르몬을 보충해서 삶의 질을 떨어뜨릴 수 있는 폐경 증후군을 효과적으로 치료하는 방법이다. 다양한 HRT 약물이 출시된 후 연구자들은 약물 사용 경험을 수집하였다(후향적 연구 방법). 그리고 HRT가 심혈관 질환 발병을 낮춘다는 잠정적 결론을 내렸다. 얼마나 획기적인가. 폐경 증후군을 치료하면 일석이조로 심혈관 질환 발병도 줄어든다! 그런데 실제로도 그럴까? 제약회사들은 많은 기대를 갖고 이중 맹검 무작위 대조군 연구를 시행했다. 실험 결과는 에스트로겐과 프로게스테론 복합 제제를 사용하면 심혈관 질환 발병이 오히려 증가할 수도 있는 것으로 드러났다. 하지만 그 정도는 미약했다. 따라서 HRT가 심혈관 질환 발병을 낮추지는 않았으나 실제로 약물을 처방하는 데는 문제가 없었다.

이중 맹검 무작위 대조군 연구는 임상시험 환경을 완벽히 통제하는 ‘실험 연구(experimental study)’라고 한다. 이에 반해 이미 수집된 빅데이터를 후향적으로(retrospective) 살펴봐서 증거를 찾는 연구를 ‘관찰 연구(observational study)’라고 한다. 아래 표는 관찰 연구에서 발견된 사실이 이후에 실험 연구를 했을 때는 입증되지 않은 예시다.

[표 1] 관찰 연구와 실험 연구의 결과가 일치하지 않은 예¹³

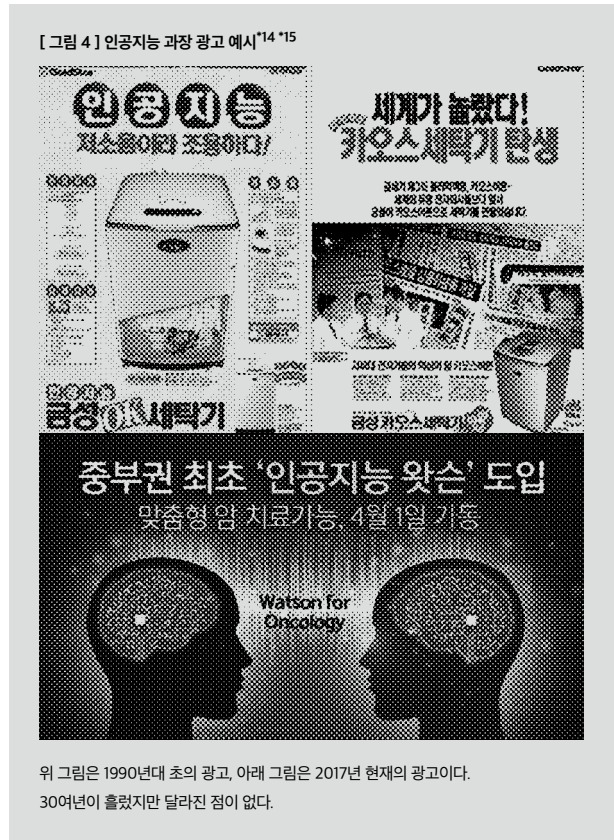
	Outcome	관찰 연구 결과	실험 연구 결과
폐경 후 호르몬 대체요법	심혈관질환 발생	감소	증가
비타민 E 고용량요법	심혈관질환 발생	감소	변화 없음
저지방 식이	심혈관질환, 암 발생	감소	변화 없음
칼슘보충요법	골절, 암 발생	감소	변화 없음
호모시스테인 감량을 위한 비타민 요법	심혈관질환 발생	감소	변화 없음

인공지능 의학연구의 속성

인공지능 기반 의학 연구에는 빅데이터가 필수적이다. 빅데이터 연구는 기본적으로 관찰 연구이다. 빅데이터 자체가 ‘이미 수집된 데이터’를 의미하기 때문이다. 바꿔 얘기하면 인공지능 기반 의학 연구는 위와 같이 관찰 연구가 갖는 한계를 고스란히 공유할 수밖에 없다. 따라서 빅데이터 연구에서 교란 요인을 적절히 통제하지 않으면 잘못된 결론을 내릴 수도 있다.

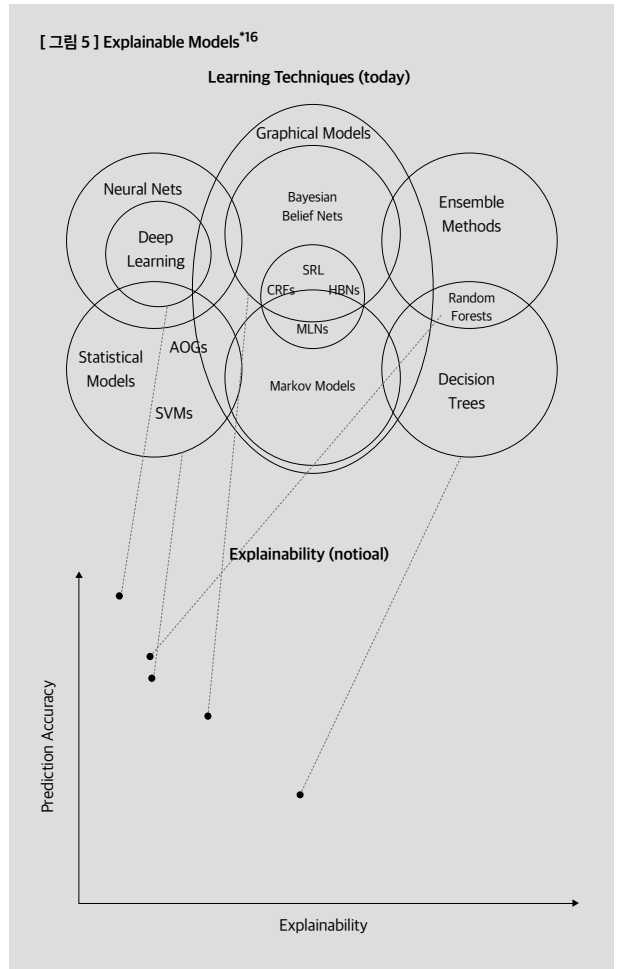
이제 인공지능 의학연구 얘기를 해보자. 다음은 가장 기본적인 질문이다.

- (1) 인공지능 알고리즘은 내적 타당도를 어떻게 검증할까?
- (2) 내적 타당도가 제대로 검증되지 않은 모델을 현실 세계에 바로 적용하는 것이 가능할까? 그렇게 하는 게 윤리적일 수 있을까?
- (3) 내적 타당도가 완벽히 검증되었다고 주장하는 인공지능 모델이 있다고 하자. 그렇다면 외적 타당도는 어떻게 검증해야 할까? 위 질문에 대한 합의된 기준은 아직 없다. 그래서 내적 타당도가 입증되지도 않은 인공지능 모델이 버젓이 ‘인공지능 모델, 의료의 혁명’이라는 과장 광고를 달고 출시되기도 한다. 예를 들어 [그림 4]를 보면 위는 1990년대 초의 광고, 아래는 2017년의 광고이다. 30여 년이 흘렀지만 달라진 점이 없다. 그렇기 때문에 의사가 문지기(gate-keeper)로서의 역할을 제대로 수행하지 못하는 순간 잘못된 인공지능 모델로 피해를 보는 환자도 나오게 될 것이다.



인공지능 모델의 ‘내적 타당도’에 대한 이야기를 해보자. 알파고가 출현한 후 ‘딥러닝’이 뜨고 있다. 딥러닝을 사용하면 내적 타당도를 높일 수 있는가? 다르게 질문을 해보자. 인공지능 알고리즘은 의학연구의 증거 강도 피라미드(pyramid of evidence-based medicine)처럼 ‘증거의 강도’를 기준으로 일렬로 세울 수 있는가? 만약 그렇다면 ‘딥러닝’은 증거 강도의 피라미드에서 가장 위에 위치할 수 있는가? 인공지능 연구에서 내적 타당도의 기준은 무엇일까? 보통 최종 결과로 제시되는 알고리즘의 성능 지표만으로는 내적 타당도를 평가하는 것에 한계가 있지 않을까?

아래 그림은 인공지능 알고리즘별로 정확성(accuracy)과 설명력(explainability)이 어떤 관계를 갖는지 보여 준다. 두 지표는 교환 관계(trade-off)를 갖는다. 설명력이 높은 모델은 정확성이 떨어지는 경향이 있고, 반대로 설명력이 떨어지면 정확성은 올라가는 경향이 있다.



인공지능 알고리즘은 증거의 강도에 따라 일률적으로 평가할 수 없다. 연구 목적에 따라 가장 적합한 학습 방법을 선택한다. 하지만, 인공지능 연구 논문이 보통 ‘강력한 근거’로 보여주는 연구 결과는 알고리즘의 성능지표다. 성능지표가 높게 나오면 내적 타당도가

확보된 것일까? 학습 데이터가 잘못되었다면? 검증을 제대로 못했다면? 학습 데이터에 과적합(overfitting)* 된 모델이라면?

* 과적합(overfitting)

너무 정교해서 융통성 없는 모델을 도출한 경우를 말한다. 예컨대, 신발을 만드는 상황을 가정해보자. 신발공은 10명의 발에 대한 자료를 갖고 있다. 그 10명은 공교롭게 모두 발가락이 6개다. 신발공이 만약 그 10명의 발에 꼭 들어맞게(perfectly fitted) 발가락 모양까지 본뜬 무중 양말 모양의 신발을 만든다면 발가락이 5개인 일반적인 사람들은 그 신발을 신지 못할 것이다. 따라서 보통 기성화를 만들 때는 발 볼의 너비와 발의 길이만을 반영한 융통성 있는 모델을 만들게 된다. 인공지능 모델을 개발할 때도 과적합을 하게 되면 성능지표는 월등하게 높아질 수 있지만 학습데이터 이외의 자료에는 적용할 수 없는 쓸모 없는 모형이 도출될 수 있다.

앞서 의학 연구의 속성에서 언급했듯이 의학 연구에서는 연구 결과의 내적 타당도를 검증할 수 있는 기준들이 잘 알려져 있다. 그 기준들의 본질은 객관적 증거로서 ‘설명’이 가능하다는 것이다. 인공지능 의학 연구는 이런 관점에서 내적 타당도를 검증할 수 있는 장치가 부족하며, 더욱이 설명력이 떨어지는 인공지능 모델을 사용하면 내적 타당도 검증은 더욱 요원한 일이 될 수 있다.

따라서 이 상황에서는 외적 타당도를 검증하는 작업이 더욱 중요해지게 된다. 설명력이 떨어지는 인공지능 모델이더라도 실제 의료 현장에서 어떤 효과를 갖고 어떤 프로세스로 활용될 수 있는지 충분히 검증된다면 유용하게 활용될 수 있을 것이다. (전통적인 의학연구에 내적 타당도가 매우 높은 무작위 대조군 연구(randomized controlled trial, RCT)으로 효과가 입증된 약물은 현실세계에서 대부분 효과가 있다. 그러나 인공지능 알고리즘은 그렇지 않을 수 있다는 뜻이다. 성능이 높게 나온 알고리즘 일지라도 내적 타당도를 전통적인 의학연구법만큼 담보할 수 없기 때문에 현실 세계에서의 검증을 더 치열하게 해야만 한다.)

다시 자동 심전도 판독 예제로 돌아가자

자동 심전도 판독은 인공지능 기술이다. 그 발전 역사를 찬찬히 살펴보면 ‘내적 타당도’를 높이는 과정과 ‘외적 타당도’를 높이는 과정이 유기적으로 병행해서 이루어진 사실을 알 수 있다.

구체적으로 어떤 교훈을 얻을 수 있는지 정리해보자.

- (1) 심전도 발전에는 전기 심장학이라는 기초 학문의 태동과 발전이 중요했다. - 인공지능 관련 기초 학문이 중요하다.
- (2) 피버그(Pipberger)나 카세레스(Caceres) 같은 연구자들의 자생적인 노력이 중요했다. - 연구자, 개발자의 창의적인 노력이 중요하다.
- (3) ECAN과 VA프로그램은 초기사용률이 저조했다. - 두 프로그램 모두 초반에는 임상 현장에서 효율성을 입증하지 못했다. 외적 타당도를 입증하지 못한 모델은 도태될 수밖에 없다.

(4) Mayo-IBM 합작사의 역할이 중요했다. - 적절한 시점에는 시장을 주도하는 기업과 의료기관의 협력이 중요하다.

(5) 시장의 필요에 의해 텔레메드(Telemed)에서 1969년 ECG 상용 서비스를 최초로 시작한 것과는 별개로 임상 의사가 그 서비스의 유용성을 인지하고 받아들이는 과정은 매우 더뎠다. - 기술 개발 초기부터 신기술의 임상적 유용성과 활용가능성을 타진하는 협력 작업이 필요하다.

(6) Telehealth 기반의 원격 판독 서비스는 이미 있었다. - 인공지능기반 서비스 모델은 의료서비스 수준이 열악한 원격지에서 수행한 검사를 일차적으로 선별하는 데 유용하게 활용될 수 있다.

(7) 정량적 심전도 분석을 위한 공통 표준처럼 개발된 신기술의 내적 타당도를 검증하기 위한 도구가 필요하다. - 인공지능 연구자들이 마음 놓고 쓸 수 있는 표준 데이터셋이 있어야 한다.

(8) 의료인의 역할이 여전히 중요하다. - 도메인 지식¹⁷⁾을 제공하는 것 뿐만 아니라 문지기로서 신기술의 유용성을 평가하는 중요한 역할을 담당해야 한다. 골목 대장으로 이득을 지키기 위한 제스처를 취하기보다는 철저한 전문가주의를 기반으로 선의의 피해자가 생기는 것을 방지하면서 기술과 의료의 공생적 발전을 도모하는 문지기가 되어야 한다. 미국의 정밀 의료 프로젝트(All of US)의 라이프 로그 수집을 담당하는 스크립스(Scripps) 연구소의 수장인 에릭 토폴(Eric Topol)은 미국의사협회지(journal of the american medical association, JAMA)에 실린 종설에서 인공지능 기술이 영상의학, 병리학 분야에 적극적으로 도입이 되면, 단순 반복적인 작업은 인공지능 기술에 이양되고 영상의학과, 병리과 의사는 정보 전문가(information specialist)로서 복잡한 임상 케이스를 관리하게 될 것이라 전망하였다. 바로 그 정보 전문가가 인공지능 시대에 의사가 수행해야 하는 문지기 역할의 일레다.

(9) 사용 증례의 지속적 축적이 중요하다. - 인공지능기술 사용 증례(use cases)를 지속적으로 축적하여 인공지능이 활용되는 구체적인 프로세스에 대한 정립이 필요하다.

결국 제일 중요한 것은 인공지능 연구자가 마음껏 사용할 수 있는 표준 빅데이터셋을 확보하는 것과 개발된 인공지능 기술을 지속적으로 검증하고 발전시킬 수 있는 플랫폼을 만드는 것이다.

다음 글에서는 의학 연구에서 ‘내적 타당도’를 평가할 때 왜 설명력이 중요한 지를 임상 의사 결정 과정의 특징을 바탕으로 살펴보도록 하겠다. 그리고 상기 심전도의 예제처럼 신의료기술이 내적 타당도를 지속적으로 높이면서 외적 타당도를 확보하려면 어떤 연구 환경이 필요한지에 대해서도 알아보도록 하겠다.

¹ 참고 | 급성 복증: 복강 내 장기의 염증, 천공, 폐색, 경색, 파열에 의한 복통을 수반하는, 갑자기 발병하는 복부의 상태로 보통 긴급한 외과적 처치를 필요로 하는 질환을 말함 ² 참고 | 심근경색증 ³ 참고 | 위키피디아 ‘Electrocardiogram’, <https://ko.wikipedia.org/wiki/%EC%8B%AC%EC%A0%84%EB%8F%84> ⁴ 참고 | <http://www.ems12lead.com/2009/02/20/76-yom-cc-shortness-of-breath-while-walking/> ⁵ 참고 | 음성예측도: 검사가 음성(정상)으로 나왔을 때 실제 병이 없을 확률 ⁶ 참고 | 논문 | Guglin, M. E., & Thatai, D. (2006). Common errors in computer electrocardiogram interpretation. *Int J Cardiol*, 106(2), 232-237. <http://doi.org/10.1016/j.ijcard.2005.02.007> ⁷ 참고 | 양성예측도: 검사가 양성(비정상)으로 나왔을 때 실제 병이 있을 확률 ⁸ 참고 | 방실전도차단 : 심장은 2심방 2심실로 구성된다. 동결질에서 생성된 전기자극은 심방에서 심실로 전도되어야 하는데 병적인 불응 상태나 절단에 의해 심방 자극이 심실로 전도되는 것이 지연되거나 차단되는 것을 방실전도차단이라고 한다. ⁹ 참고 | 위키피디아 ‘Automated ECG interpretation’ ¹⁰ 참고 | https://en.wikipedia.org/wiki/Automated_ECG_interpretation ¹¹ 논문 | Pentti et al.(2016). Eyewitness to history: landmarks in the development of computerized electrocardiography, *Journal of Electrocardiology* 49 (2016) 1 - 6 ¹² 참고 | 교란 요인: 보고자 하는 실험결과에 영향을 미치는 요인이다. 아스피린 예시에서는 ‘물 많이 먹기’ ‘옷 벗어 냉각하기’ 등이 해당된다. ¹³ 참고 | 정세영 역. (2017). 직관으로 이해하는 의학 통계학. 대한의학서적 ¹⁴ 참고 | <https://www.kyuh.ac.kr/main.asp?ver=pc> ¹⁵ 참고 | <http://v.media.daum.net/v/20171225090028486> ¹⁶ 참고 | DARPA. (2016). Explainable Artificial Intelligence.(pp.9) ¹⁷ 참고 | 도메인 지식이란 의학 교과서에 나온 객관화할 수 있는 내용뿐만 아니라 임상현장에서 체득한 다양한 경험을 포함한다.

헬스케어 빅데이터 딜레마와 해결 방안

‘헬스케어’와 전통적인 ‘의료 서비스’는 그 개념 간의 차이가 있다.

전통적인 의료 서비스가 환자의 질병 치료에 방점을 둔다면, 헬스케어는 ‘정보통신기술’(ICBM: 사물인터넷, 클라우드, 빅데이터, 모바일)을 기반으로 하여 환자 뿐 아니라 일반인의 질병을 예측하고 일상을 관리함으로써 각 개인의 건강 수명을 연장하는 데 관심을 둔다는 점이 다르다.

헬스케어 빅데이터에 집중된 시선 : 2차적 활용의 딜레마

질병을 잘 예측하고 관리하는 데에는 ‘헬스케어 빅데이터’라는 연료가 필요하다. 고혈압 환자의 예를 들어 보자. WHO에 따르면, 전 세계의 고혈압 환자는 약 50억 명으로 추산된다. 전통적인 임상 현장에서는 고혈압 신약을 개발하기 위해 상당한 시간과 비용(실패 비용 포함 4~11조 원 추산)을 지불해야 했고 연구에 참여하는 환자는 아직 검증되지 않은 시험약을 투여 받아야 하는 위험이 있었다. 그러나 인공지능으로 사용자별 고혈압이 발생하는 시점을 예측하고 실시간으로 복용 약물을 체크해 주거나 음주 가능 여부 등을 알려주는 시스템이 있다면 50억여 명의 고혈압 환자가 발생하기 전 이를 예방하는 데 큰 도움이 될 것이다. 이 때 인공지능 분석에는 이미 수집된 고혈압 관련 헬스케어 빅데이터가 기반이 된다. 이처럼 헬스케어 빅데이터를 기반으로 한 새로운 건강 관리 패러다임은 의료 비용을 절감하는 동시에 의료의 질을 향상시킬 수 있다는 것이다.

그렇다면 헬스케어 빅데이터는 무엇을 의미하는가?

헬스케어 빅데이터는 아래 표와 같이 정의되고 분류될 수 있다.

[표 1] 헬스케어 빅데이터 종류 구분	
구분	종류
진료 데이터	전자의무기록, 전자건강기록 정보, 처방정보, 입/퇴원 기록, 의료영상자료 등 데이터
임상연구 데이터	의약품(drug) 임상시험 데이터, 의료기기(device) 임상시험 데이터, 유전자 연구 데이터, 인체유래물 연구 데이터, 조사관찰 연구 데이터, 개인정보를 직·간접적으로 활용한 연구 데이터
공공기관 데이터	자격 및 보험료 관련 데이터, 진료 내역, 건강검진 결과, (건강보험공단, 통계청, 사망정보 등 공공기관에서 수집·보관·관리하는 데이터 건강보험심사평가원 등)
기기 기반 데이터	의료기기 및 환자 모니터링 장치 기반 데이터
오믹스(omics) 데이터	유전체(Genome), 전사체(Transcriptome), 단백질체(Proteome), 대사체(Metabolome), 후성유전체(Epigenome), 지질체(Lipidome) 등 다양한 분자 수준에서 생성된 여러 데이터
라이프로그(lifelog) 데이터	대체로 웨어러블, 홈 모니터링 장치, IoT, 모바일 앱 등에서 수집된 데이터로, 한 개인의 일상생활 활동에 관한 모든 데이터(체중, 심박수, 혈당, 몸무게, 식습관, 운동 습관, 약물 복용 여부, 행동과 정서에 대한 데이터 등으로 많은 사람들이 스스로 건강정보를 관리하기 시작함)
앱·소셜 미디어 데이터	건강 포털, 의사 포털, SNS에서 수집된 다양한 헬스케어 관련 데이터

헬스케어 데이터는 전통적으로는 병원 내에서 생성되고

축적되었지만, ICBM의 발전에 따라 라이프로그 데이터, SNS 데이터와 같은 병원 밖의 새로운 헬스케어 데이터들이 생성되어 헬스케어 빅데이터 개념 내에 새롭게 포함되었다. 예를 들어, 페이션츠트라이크미(Patientslikeme)는 2004년 루게릭 환자들의

치료를법 찾기 위해 MIT 출신의 3명의 엔지니어가 모여 만든 SNS 서비스로 시작되어 현재는 미국 등을 포함한 전 세계 25만 명의 중증 질환 환자들이 모인 커뮤니티로 성장하였다. 이 SNS에서 환자들은 증세 진행 과정, 약의 부작용과 효능, 질병의 재발 및 관리 등 본인의 질병에 관한 기록을 남기고 노하우를 공유한다.

또 축적된 데이터는 전문적으로 분석되고 환자들에게 제공되어 환자들이 해당 질병을 이해하고 예측하는데 도움을 주기도 한다. 전통적인 의료 시스템 내에서는 시판된 약품 또는 임상시험 중인 약물을 추적 관찰하기 위해 상당한 시간과 비용이 든다. 그러나 환자들로부터 제공 받는 해당 질환과 관련한 여러 정보와 실시간 피드백은 ‘real world data’로 여러 의료 관계자(의료 기관, 제약회사 등)가 이 데이터를 빠르게 분석하여 효율적으로 질병을 연구하거나 신속하게 치료제를 개발하는 데 큰 도움을 줄 수 있다. 실제로, 글로벌 제약회사인 노바티스(Novartis) 등은 SNS 헬스케어 빅데이터를 활용하여 신약을 개발 중에 있다.

특정인의 질환을 정확하게 예측하는 알고리즘을 개발하기 위해서는 특정인의 정보도 필요하지만, 특정인과 동일한 질환을 갖고 있거나 유사한 질환을 지니고 있거나 지닌 경험이 있는 환자들의 ‘헬스케어 빅데이터’를 활용하는 것이 절대적으로 중요하다. 헬스케어 데이터가 많이 주어지고 또 데이터끼리 연계될수록 개인에게 맞춤형된 헬스케어 서비스를 보다 많이 제공받을 수 있기 때문이다.

여기서 하나의 딜레마가 있다. 헬스케어 빅데이터 처리에 대한 정보 주체로부터의 동의 문제가 발생하기 때문이다. ‘헬스케어 빅데이터’의 대부분을 차지하는 건강 정보, 유전 정보는 개인정보에 대한 기본법인 「개인정보보호법」에 따라 ‘민감정보’에 해당된다.

「개인정보보호법」 제23조(민감정보의 처리 제한)
① 개인정보처리자는 사상·신념, 노동조합·정당의 가입·탈퇴, 정치적 견해, 건강, 성생활 등에 관한 정보, 그 밖에 정보주체의 사생활을 현저히 침해할 우려가 있는 개인정보로서 대통령령으로 정하는 정보(이하, “민감정보”라 한다)를 처리해서는 아니된다. 다만, 다만 각 호의 어느 하나에 해당하는 경우에는 그러지 아니한다.
1. 정보주체에게 제15조제2항 각 호 또는 제17조 제2항 각 호의 사항을 알리고 다른 개인정보의 처리에 대한 동의와 별도로 동의를 받은 경우
2. 법령에서 민감정보의 처리를 요구하거나 허용하는 경우
「개인정보보호법」 시행령 제18조(민감정보의 범위) (중략) 1. 유전자검사 등의 결과로 얻어진 유전정보 (중략)

‘건강정보’에는 키, 몸무게, 진료기록 등 다양한 층위의 정보가 포함될 수 있으나, 「개인정보보호법」은 이를 구분하지 않고 건강 정보로 규정한다. 건강정보로 통칭되는 여러 정보들은 이 법에 따라 다른 정보보다 민감한 정보로 분류되며, 건강정보를 ‘처리’하려는 경우 법령에서 민감 정보의 처리를 요구하거나 허용하는 경우가 아니라면 정보주체로부터 별도의 동의를 받아야 한다.

글 | 유소영 mesoyoung@gmail.com

이화여자대학교에서 철학과 심리학을 공부하였으며, 미국 Boston College에서 심리학으로 석사 학위를, 이화여대에서 생명윤리정책으로 박사 학위를 받았다. 현재는 서울아산병원 헬스이노베이션 빅데이터센터와 임상연구보호센터에서 정책 기획을 담당하고 있다. 신기술 발전에 따른 인간의 위치와 역할의 변화에 관심이 많으며, 주 연구 분야로는 ‘ICT 헬스케어 정책’, ‘데이터 민주주의’, ‘신경인문학(Neurohumanity)’ 등이 있다. 또한, 대통령 직속 4차 산업혁명 위원회 헬스케어 특별 위원회, 대한기관윤리심의기구협의회(KAIRB), 복지부 공용 기관생명윤리위원회(IRB) 운영 위원회 등에서 위원으로 활동하며, 헬스케어 정책 마련을 위해 노력하고 있다.

<p>「개인정보보호법」 제2조(정의)</p> <p>2. “처리”란 개인정보의 수집, 생성, 연계, 연동, 기록, 저장, 보유, 가공, 편집, 검색, 출력, 정정(訂正), 복구, 이용, 제공, 공개, 파기(破棄), 그 밖에 이와 유사한 행위를 말한다.</p> <p>따라서, 건강정보를 수집하는 것 뿐만 아니라 이미 수집된 진료정보, 처방정보 등을 가공하거나 다른 데이터와 연계하기 위해서도 원칙적으로는 해당 목적에 따라 동의를 받아야 한다.</p>
--

또한, 빅데이터를 구축하기 위한 핵심은 흩어져 있는 여러 헬스케어 데이터(건강정보 등)를 종단적(longitudinal)으로 연계·결합하는 것이다. 즉, 헬스케어 빅데이터는 다양한 목적으로 이미 축적된 헬스케어 데이터를 2차적 용도”로 활용할 수 있을 때, 그리고 고유식별정보(주민등록번호, 외국인등록번호 등)로 한 개인의 다양한 데이터를 연계하여 개인별 건강상태에 대한 정확한 흐름을 분석할 수 있을 때 그 가치가 발현된다.

예컨대 심전도 기계만 해도 홀로 초당 1,000번의 측정이 이루어지는데, 이 데이터 중 극히 일부만 이용되고 나머지는 대부분은 버려진다. 그 버려진 부분이 오히려 환자의 상태나 진료, 질병 예측 등을 알려줄 수 있는 정보일 수도 있는데 말이다. 데이터를 수집, 저장, 연계, 분석하는 데 기술적, 비용적 어려움이 있던 시절에는 데이터를 버리는 것이 더 효율적이었겠지만 정보통신기술이 보편화된 현재는 그렇지 않다. 이 데이터가 이미 수집된 다른 데이터와 고유식별정보를 통해 종단적으로 연계될 수 있거나 다른 환자들의 데이터와 함께 분석될 수 있다면, 심전도 측정의 목적 외의 다른 2차적 목적(예, 심전도 연구, 심전도 기계 개발 등)으로 유용하게 활용될 수 있을 것이다.

그런데 건강정보와 마찬가지로, 고유식별정보 역시 「개인정보보호법」에 따라 다른 법령에서 특별한 규정이 있지 않다면, 구체적 고유식별정보 처리 목적에 따라 모든 정보주체로부터 개별적으로 동의를 받는 것이 원칙이다.

<p>「개인정보보호법」 제24조(고유식별정보의 처리 제한)</p> <p>① 개인정보처리자는 다음 각 호의 경우를 제외하고는 법령에 따라 개인을 고유하게 구별하기 위하여 부여된 식별정보로서 대통령령으로 정하는 정보(이하 “고유식별정보”라 한다)를 처리할 수 없다.</p> <p>1. 정보주체에게 (중략) 다른 개인정보 처리에 대한 동의와 별도로 동의를 받은 경우</p> <p>2. 법령에서 구체적으로 고유식별정보의 처리를 요구하거나 허용하는 경우</p>

그러나 이미 수집된 고유식별정보 마저 구체적 처리 목적별로 모두 다시 동의를 받아야 한다면, 헬스케어 빅데이터는 영원히 구축될 수 없다. ‘충분한 정보에 근거한 개별 동의(informed and specific consent)’라는 정형화된 체계는 헬스케어 빅데이터 시대에 더 이상 적합하지 않다.

물론, 「개인정보보호법」 제18조(개인정보의 목적 외 이용·제공 제한)에서는 정보주체 또는 제3자의 이익을 부당하게

침해할 우려가 없으며, 통계 작성 및 학술 연구 등의 목적을 위하여 필요한 경우로서 특정 개인을 알아볼 수 없는 형태로 개인정보를 제공하는 경우일 때는 데이터를 2차적 목적으로 활용할 수 있다고 명시하고 있다. 하지만 이는 어디까지나 개인을 알아볼 수 없도록 비식별 처리된 정보에 한한 것으로, 고유식별정보를 활용하는 경우는 해당되지 않는다.

또한 헬스케어 데이터의 2차적 활용의 중요성을 인지한 일부 개별법(예, 보건의료기술진흥법, 암관리법, 희귀질환관리법 등)에서는 보건복지부 등에서 요청 시 정보주체의 별도의 동의 없이 고유식별정보를 활용하여 국가나 공공기관에서 보관하고 있는 헬스케어 정보들을 연계 및 결합할 수 있다. 그러나 민간기관의 경우 정보주체의 동의 없이 고유식별정보를 활용하여 헬스케어 데이터를 연계할 수 있는 법적 근거가 미비하다.^{2*3}

많은 사람들은 동의 제도가 빅데이터 시대에 한계로 작용한다는 것에 대해서는 충분히 공감할 수 있을 것이다. 그렇지만 나를 알아볼 수 있는 데이터가 나의 동의 없이 헬스케어라는 목적 아래 누구에게나 활용되는 것을 쉽게 반기지는 않을 것이다. 특히 나의 민감정보나 고유식별정보의 2차적 활용에 따라 내가 받는 직접적인 이익은 눈에 보이지 않고 잠재적 이익은 미래의 일이라 가정해 본다면 더더욱 그러하다. 이 딜레마(데이터 2차적 활용 필요성 VS. 개인정보의 자기결정권)는 과연 해결될 수 있을까?

딜레마 해결을 위한 세 가지 전략과 한계

이 딜레마 상황을 최소화하기 위한 많은 시도들이 있다. 현재 가장 많은 호응을 얻고 있는 세 가지 전략을 소개하고, 이 제안의 매력 포인트와 한계에 대해 언급하고자 한다.

1) 정직한 중개인(Honest Broker) : 제3자 세우기 전략

가장 전통적인 제안 중 하나가 ‘정직한 중개인(honest broker, HB)’ 전략이다. HB는 데이터의 연계 요청자나 요청자의 활용 목적과는 관련 없는 ‘독립적인 데이터 연계 관리자’이다. HB를 제안하는 많은 이들은 동의받지 않은 고유식별정보가 오직 HB에 의해서만 2차적 목적을 위해 연계·결합되고 요청자는 고유식별정보가 제거된 최종 결합물을 제공받을 수 있다면 우리가 우려했던 개인정보 보안이나 사생활 침해의 문제가 최소화될 수 있다고 말한다.

이 방법은 실제 글로벌 의료기관에서 활용되고 있다.⁴ 출판이나 발표 등을 통해 공개된 자료를 분석한 결과, 의료기관 중 HB가 있는 기관은 총 10곳이 있다. HB는 헬스케어 빅데이터 플랫폼에서 개인식별정보를 자동적으로 비식별 처리하는 System Honest Broker와 인간 중개인인 Human Honest Broker로

구분된다. 이 둘은 2차적 목적으로 동의받지 않은 헬스케어 데이터를 익명처리하거나 고유식별정보를 활용하여 연계한 후 비식별 처리한 결과물만을 데이터 요청자에게 전달함으로써 개인식별정보를 보호하면서도 효율적으로 헬스케어 데이터를 활용하는데 그 역할을 수행한다.

그러나 여전히 관련 법령상 동의 없이 고유식별정보를 2차적으로 활용할 수 있다는 명시적 조항이 있지 않은 이상, 또는 HB를 활용하여 고유식별정보를 처리할 수 있다는 법적 근거가 있지 않는 이상, 실제 HB를 도입하기에는 한계가 있다. 또한 HB를 활용할 수 있는 법적 근거가 마련 될지라도, 여전히 자기결정권 측면에서의 동의 문제나 고유식별정보 제공자가 가질 이익 측면에서는 명료한 답을 내지 못하고 있다.

실제로 2016년 6개 범부처에서 합동으로 마련한 <개인정보 비식별 조치 가이드라인>에 따라 보건복지 분야 빅데이터 분석을 위해 서로 다른 사업자가 보유하고 있는 정보집합물을 결합하는 경우 ‘사회보장정보원’이 국가 HB로 지정되었지만 위와 같은 이유로 민간 기관 간의 고유식별정보 연계 기능은 하지 못하고 있다.

2) 헬스케어 블록체인 & 메디토큰 : 화폐로 보상하기 전략

‘헬스케어 블록체인’도 이 문제를 해결하기 위한 전략 중 하나로 제안된다. 헬스케어 블록체인은 의료정보에 대한 접근 권한을 의료 공급자가 아닌 환자에게 부여하여, 환자 본인만이 궁극적으로 자신의 데이터를 복호화할 수 있고 접근 권한을 본인만이 자유롭게 설정할 수 있도록 한다. 이로써, 정보주체의 데이터에 대한 자기결정권을 최대화하면서도 정보의 투명성 및 보안성을 담보할 수 있다는 것이다.

<p>MEDI BLOC WHITEPAPER⁵, 1. 서론 일부 발췌</p> <p>현재 의료정보시스템은 의료기관 중심으로 운영되고 있다. 그리고 의료기관 밖으로 의료 정보를 공유하는 일은 개인 정보 보호를 위해 환자 본인의 자신의 의료기록을 요청하는 경우를 제외하고 허용되지 않는다. 이러한 의료기관 중심의 의료정보 관리체계는 개인의 의료 데이터를 여러 병원에 분산시켰고, 파편화된 의료 데이터로 의료 서비스의 질을 저하시켰다. 의료 연구나 시를 위한 의료정보에 대한 요구 역시 날로 증가하고 있으나 데이터의 공급은 턱없이 부족한 상황이며, 현재 시스템에서는 데이터의 신뢰성도 충분히 담보하기 힘들다. (중략) 여기 관에 흩어져 있는 의료정보 뿐만 아니라 스마트폰을 포함한 여러 기기를 통해 생산되는 모든 의료정보를 안전하게 통합하여 관리할 수 있게 하는 의료정보 플랫폼을 제공한다. (중략) 의료정보 소유권 및 관리권한을 재분배하고 이를 기반으로 의료 전반에 걸친 혁신적인 변화를 만들어낼 수 있을 것으로 확신한다.</p>
--

특별히 헬스케어 블록체인 플랫폼에서는 메디토큰(Medi Token)을 발행하여 이를 중심으로 플랫폼 내 경제 생태계를 구축할 계획을 가지고 있다. 이 생태계에 기여하는 참여자들은 그 기여도에 따라 메디토큰을 보상받게 되는 데, 의료 소비자뿐 아니라 의료정보의

생산에 기여한 의료 공급자도 기여 정도에 따라 정당한 보상을 받을 수 있을 것이라 설명한다. 예컨대, 폐암 환자들이나 의료 기관이 폐암 정보를 원하는 제약회사나 연구자에게 폐암 데이터를 제공할 경우, 이에 대한 제공의 정도가 많을수록 많은 금전적 혜택을 받을 수 있다.

<p>MEDI BLOC WHITEPAPER, 3. 메디블록 기술적 세부사항 일부 발췌</p> <p>(중략) 타인의 데이터를 얻고자 하는 경우 메디블록 실시간 검색 시스템을 통해 데이터를 찾는 것은 물론, 찾고자 하는 데이터의 조건과 데이터 제공에 대한 보상 요건 등을 명시해 메디블록 네트워크에 알릴 수 있다. 개별 사용자는 본인의 데이터가 이 조건에 부합하는지 여부를 개인 기기에서 판별한 후 푸시 알림 기능등을 통해 데이터 거래에 참여할 수 있다. 이 모든 기능은 사용자의 능동적인 참여 없이도 수행이 가능하도록 백그라운드에서 이루어진다.</p>
--

그런데 헬스케어 빅데이터를 구축하기 위해서는 무엇보다 데이터 공유(data sharing)가 필요하다. 그러나 우리가 최근 비트코인 사태에서 유추할 수 있듯이, 만일 더 많은 보상을 받기 위해 정보 주체가 원하는 가격이 제시될 때까지 본인의 데이터를 제공하지 않는다면 어떻게 될까? 이 경우 정보 사용에 대한 자기결정권은 최대 발현될 수 있겠지만 데이터의 공유 측면에서는 상당히 위험하다.

3) 규제 샌드박스 : ‘달이나 달걀이나’의 해소 전략

마지막은 ‘규제 샌드박스(regulatory sandbox)’ 전략이다. 규제 샌드박스란 어린이들이 자유롭게 노는 모래 놀이터처럼, 제한된 환경(제한된 주제, 제한된 프로젝트)에서 한시적으로 규제를 풀어(탄력 적용), 신산업을 테스트(시범 사업)하도록 하는 것을 의미한다. 현 정부는 지난 2017년 11월, 신산업 분야에 규제 샌드박스를 도입하겠다고 발표하여 큰 관심을 받기도 하였다.

특별히 헬스케어 분야에서 규제 샌드박스를 받기는 이유가 있다. 헬스케어 분야는 인간의 건강을 다루는 분야로 다른 어떤 분야보다 많은 규제가 부과되어 있다. 문제는 규제가 아직 마련되지 않은 상태에서 혁신적인 헬스케어 산업이 기하급수적인 속도로 발전하고 있다는 점이다. 현재와 같은 상황에서 혁신 제품을 위해 관련 현행 법령을 모두 개정하거나 규제를 마련할 때까지 기다리는 것은 헬스케어 신산업을 성장할 수 있는 골든타임을 놓치게 할 뿐만 아니라 해당 신기술로 혜택을 받을 수 있는 환자에게도 이롭지 못한 결과를 가져온다. 그렇기 때문에 정부에 의한 강력한 탑-다운 식의 선 규제 전략은 더 이상 적합하지 않다는 지적이 힘을 얻고 있다.

만일 헬스케어 규제 샌드박스에서 일부 위험이 낮은 헬스케어 산업 분야에 대해 빅데이터 활용에 대한 규제를 일정 기간 면제 또는 유예함으로써 데이터를 활용한 신 헬스케어 상품이 빠르게 출시되고 이에 대한 안전성(safety), 안정성(stability)을 인정받을 수 있다면, 수요자와 공급자에게 모두 이익이 될 수 있다. 이후 이 시범 운영에서 나온 문제점들을 분석하여 사후 모니터링,

사후 규제를 마련하는 형식으로 규제 방식을 변화시키는 것이 규제 샌드박스 도입의 목적이다.

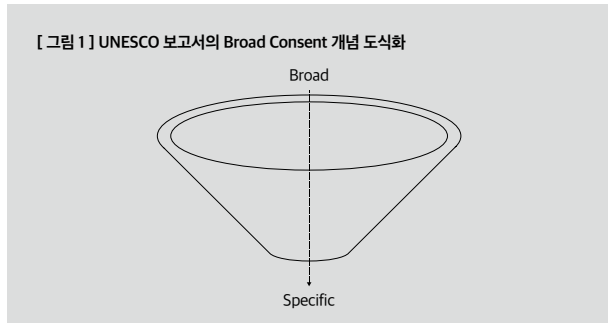
규제 샌드박스는 님이나 달걀이나 하는 문제(헬스케어 빅데이터를 2차적 활용할 수 있어야 신기술이 발전할 수 있다 VS. 신기술이 대중에게 주는 이익이 있어야 헬스케어 빅데이터의 2차적 활용을 허용할 것이다)의 문제의 답을 찾을 수도 있다. 규제 샌드박스는 일부 프로젝트에서 빅데이터의 2차적 활용이 일시적으로나마 가능하도록 하는 동시에 이를 통해 어떤 규제 정책을 가지고 갈지를 결정한다는 측면에서는 고무적이다. 그러나 전체 헬스케어 빅데이터 산업에서 규제 샌드박스가 보편적으로 적용되기란 어렵다.

달레마 해결을 위한 프레임 전환 : UNESCO <Report on Big Data and Health>

이런 여러 고민 속에 2017년 9월, 유네스코(UNESCO) 국제생명윤리위원회(International Bioethics Committee, IBC)에서 <Report on Big Data and Health>⁶라는 흥미로운 보고서를 발표했다. 이 보고서는 우리에게 몇 가지 프레임 전환을 요청한다.

첫째, ‘포괄 동의(broad consent)’에 대한 인식 변화이다. 포괄 동의는 동의하는 현재 시점에서는 구체적이지 않지만 동의한 카테고리 하에 추후 여러 목적으로 활용될 수 있다는 것에 동의하는 것을 의미한다. 포괄 동의는 데이터 활용 목적이 발생할 때마다 해당 목적에 따른 상세 항목(개인정보의 수집 및 이용 목적, 수집하려는 개인정보 항목, 개인정보의 보유 및 이용 기간 등)에 대해 충분히 설명하고 동의를 받아야 하는 ‘구체적 동의(specific consent)’ 또는 ‘충분한 설명에 근거한 동의(informed consent)’와는 그 동의의 질(quality)이 다른 것으로 인식되어, 헬스케어 분야 법령에서는 보편적으로 포괄 동의를 인정하지 않고 있었다.

그런데 이 보고서는 특별히 보건 의료 연구 분야에서, ‘포괄 동의’가 ‘충분한 설명에 근거한 동의’ 또는 ‘구체적 동의’와 반대되는 개념이 아닌, 본인의 정보로 수행 가능한 범위의 연구에 동의한 것이라고 지적한다.



즉, 전통적인 동의 모델인 구체적 동의 또는 충분한 설명에 근거한 개별 동의만이 정보의 자기결정권을 보장하는 것이 아니며, 포괄적으로 범주에 대한 동의를 하여 다양한 활용을 가능하도록 한 것 역시 정보주체의 정보 자기결정권의 발현이라는 것이다. 이렇게 IBC는 헬스케어 빅데이터의 최대 활용 가능성을 검토하는 한편, 동시에 인간의 기본권 역시 존중할 것을 제안하고 있다.

둘째, 이 보고서에서는 ‘빅데이터 제공에 대한 보상과 데이터의 소유권’에 대해서 이야기한다. 빅데이터의 핵심인 데이터의 실시간 분석, 연계, 대규모 데이터 베이스의 공유 등의 특성으로 인해 더 이상 데이터의 ‘소유권’의 개념으로는 윤리적·법적 문제를 해결할 수 없다고 주장한다. 그렇기 때문에 헬스케어 빅데이터의 특징에 따른 새로운 대안적 규범 체계가 필요하다고 제안한다. IBC는 이 시대에서의 빅데이터는 인류의 공공선 관점으로 보아 누구나 사용 가능한 인프라를 구축하고 시민이 공공선을 위해 자신의 데이터를 공유하거나 기부할 수 있는 연대감을 형성할 것을 요청한다. 또한 헬스케어 데이터의 관리자 입장에서는 빅데이터에 기여한 집단과 개인에게 주는 혜택을 고민하도록 하여 데이터 소유권의 관점으로부터 관리자의 책무와 혜택 공유로의 프레임 전환을 촉구하고 있다.

앞으로 가야할 길

여전히 남아 있는 질문이 있다. 헬스케어 빅데이터의 2차적 활용을 위해 프레임을 전환하거나 새로운 규범을 마련하는 것도 중요하지만, 우리 사회가 추구해야 하는 가치가 무엇인지, 또 그것이 왜 필요한지 그 근거를 마련하는 것이 선결되어야 할 중요한 문제다.

<제2의 기계 시대>⁷ 중 발췌
 미래의 성공은, '기술적인 면에서 어떤 선택을 하는가' 뿐만 아니라, 새로운 조직과 제도의 공동 발명에 달려 있다. 우리 세대는 역사상 그 어떤 세대 보다도 세상을 바꿀 기회를 더 많이 물려받았다. (중략)
 [기술의 발전에 따라] 우리가 할 수 있는 일을 제약하는 요인들이 줄어들수록 필연적으로 우리가 무엇에 가치를 두는지가 점점 중요하다.

1) 우리는 IBC가 제안한 바와 같이 사회를 위해 자신의 데이터를 공유하는 공리적인 선택을 할 수 있을까?

이에 대한 답변은 공공선을 위해 자신의 데이터를 공유하거나 기부하게 되는 동력이 될 것으로 보인다.

만일 우리가 공공의 선을 위해 데이터의 공유에 동참할 수 없다고 판단한다면, 이를 거부하는 이유가 무엇일까? 참여를 결정하는 이유보다 거부를 결정한 원인을 아는 것이 문제를 해결하는 데 도움이 될 것이다.

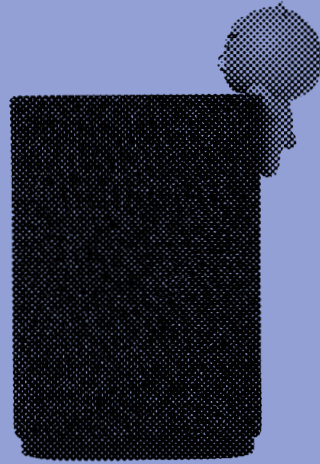
2) 포괄 동의가 우리 사회에 일반화되었다고 가정하였을 때, 데이터가 정보주체가 동의한 범주 내에서 활용된다는 것을 어떻게 보장받을 수 있을까?

포괄 동의 내 정보 활용의 적정성을 담보할 수 있도록 독립적인 검토 기관 또는 위원회의 고려가 필요할 것이다.

이 때, 독립적인 기관 또는 위원회는 정보가 관리되지 않아 정보 공유지의 비극이 되지 않도록 추적 검토하며 활용에 승인하는 역할을 담당하면서도 한편으로는 빅데이터의 집합소가 빅브라더가 되지 않도록 고민해야 한다. 이 경우 빅데이터 시대에 포괄동의 체계로의 전환은 동의권의 축소가 아닌, 정보주체가 결정한 범주 내의 적정성을 다시 한번 검토받았다는 측면에서 오히려 자기결정권의 연장과 확대로 평가받을 수 있다.

⁶1 참고 | 데이터의 2차적 사용이란 동의 받은 목적 외의 다른 목적으로 데이터를 활용하는 것 또는 동의 받은 목적 외의 제3자에게 해당 데이터를 제공하여 동의 받은 목적 외의 다른 목적으로 데이터를 활용하도록 하는 것을 말한다. ²2 정책 보고서 | 권순영, 유소영 등. (2016). 개인정보의 연구 목적 처리를 위한 법, 제도 개선방안 연구, 개인정보보호위원회 2016: 98-135. ³3 참고 | 정부 또는 국가 기관 간의 고유식별정보를 활용한 데이터 간의 연계를 허용한 것은 해당 기관이 지니고 있는 "공적인" 목적을 고려했기 때문이라고 생각한다. 그러나 "공적" 목적 이 늘 항상 공공기관에서만 발현될 수 있는 것이 아니다. 따라서 민간기관이 공적 목적으로 헬스케어 데이터 연계를 필요한 경우에도 이가 허용되지 않는다는 보다 타당한 명분이 있지 않는 한 해당 예외 조항에 대한 논란은 지속적으로 발생할 것이다. ⁴4 논문 | Choi, H.J., et al. (2015). Establishing the role of honest broker: bridging the gap between protecting personal health data and clinical research efficiency, PeerJ, 3, e1506. ⁵5 참고 | https://medibloc.org/whitepaper/medibloc_whitepaper_kr.pdf ⁶6 참고 | <http://unesdoc.unesco.org/images/0024/002487/248724E.pdf> ⁷7 참고 | 에릭 브린올프슨, 앤드루 맥아피. (2014). 제2의 기계시대. 청림출판

카카오미니에게 궁금한 것



카카오미니를 깨우는 “헤이, 카카오!”. 이 호출 명령어(wake-up word)와 카카오미니가 호출 명령어를 인식하는 원리 및 과정에 대해 알아보고, 카카오미니를 더욱 편리하게 사용하기 위해 진행 중인 음성 인터페이스 AI 기술에 대해 설명하는 내용을 첫 번째 글로 준비했습니다. 아울러 카카오미니가 사람의 언어를 이해하고 진화하는 방법, 카카오미니 서비스가 만들어지는 과정, 그리고 이 모든 것을 가능하게 해주는 개발 플랫폼인 ‘카카오(아이) 오픈 빌더(Kakao I Open Builder)’에 대한 설명을 이번 장의 두 번째 글로 담았습니다.

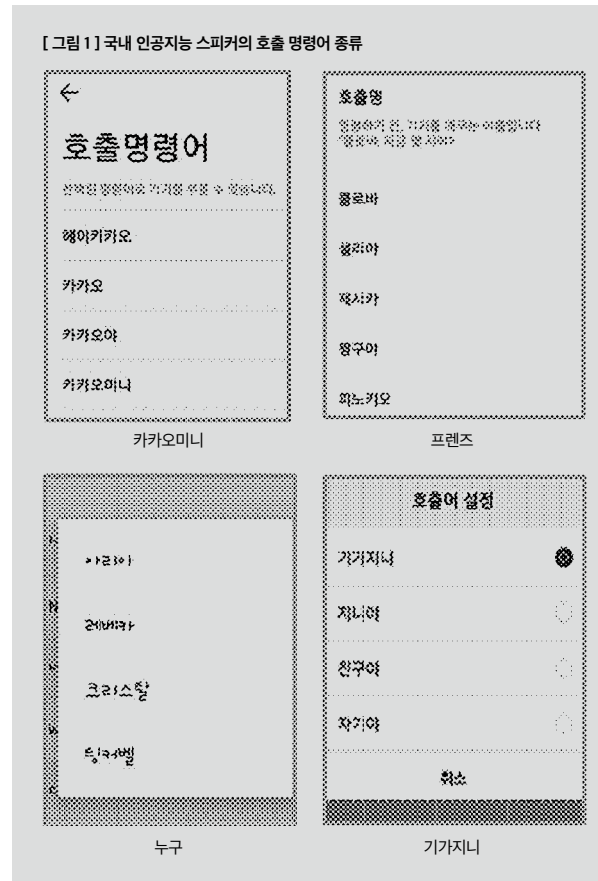
“헤이, 카카오!”를 불러야 하는 이유

카카오미니를 켜면, “이제 ‘헤이, 카카오!’라고 불러주세요”라는 인사말을 한다. 어느 음성 인식 스피커와 마찬가지로, 카카오미니에게 원하는 명령을 하려면 일단은 무언가 불러야 카카오미니가 귀를 기울이고 사용자의 명령을 들을 준비를 한다. “헤이, 카카오! 아이유 노래 들어”와 같은 식이다.

‘헤이, 카카오!’와 같이 스마트 스피커를 부르는 단어를 호출 명령어(wake-up word) 또는 호출어라고 한다. 그냥 “야, 노래 좀 들어봐”라고 하면 좋겠지만, 그러면 스피커는 이용자가 자신을 부르는 지를 인식하지 못한다. 조금 귀찮더라도 정해진 호출 명령어를 불러줘야 한다. 각 제품마다 여러 가지 호출 명령어가 있으며 사용자는 이들 중 하나를 선택해서 사용한다.

호출 명령어는 어떻게 정해질까?

스피커마다 호출 명령어는 다르다. [그림 1]과 같이 국내에 출시된 인공지능 스피커는 그 이름도 많지만 이를 부를 때 사용하는 호출 명령어는 더 다양해서 헷갈릴 정도다.



이렇게 다양한 호출 명령어는 기술적인 측면과 사업적 판단이 함께 고려되어 선정된 결과다.

일반적으로 호출 명령어로는 3음절에서 5음절의 단어가 적당하다. 너무 짧으면 스마트스피커가 호출 명령어인지 아닌지를 식별하기 어렵고, 너무 길면 이용자가 말하기 힘들다. 단어의 끝이 /ㅏ/, /ㅑ/ 모음인 명령어는 이용자가 발음하기 쉬울 뿐만 아니라, 해당 단어의 음성은 또렷하기에 스마트스피커가 알아듣기도 용이하다. 반면에 /ㄱ/, /ㄴ/ 모음으로 끝나는 단어는 음성의 크기가 작아서 적절하지 않다. 또한 /ㅋ/, /ㅌ/, /ㅍ/, /ㅍ/ 과 같이 구분이 잘 되는 발음을 만들어 내는 단어가 호출 명령어로 적합하다.

이러한 기술적 특성과 함께 기업명이나 플랫폼 명칭 등을 포함하거나 브랜드 아이덴티티에 대한 고려가 종합적으로 이뤄져, 호출 명령어가 선정된다. 결국 사용자가 부르기 쉽고 회사의 이미지에 어울리는 단어가 선정되는 것이다.

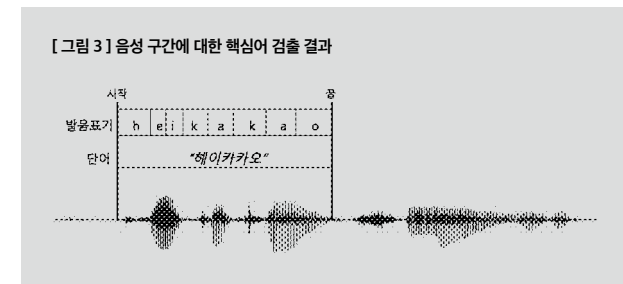
호출 명령어는 단순히 스피커를 부르는 용도로만 쓰이는 게 아니라 그 자체로 브랜드가 되어가고 있다. 아마존의

알렉사(Alexa)는 완전히 플랫폼 명칭이 되었고, 구글 또한 최근 구글 어시스턴트가 아닌 헤이 구글(Hey Google)이란 호출 명령어를 스마트 스피커의 브랜딩에 적극 활용하고 있다. 호출 명령어가 상품명보다 더 기억하기 쉽고 사람들에게 친숙하기 때문이다.



카카오미니는 어떻게 “헤이, 카카오!”를 인식할까?

그렇다면 카카오미니는 일상 대화와 호출 명령어를 어떻게 구분할까? 스피커가 호출 명령어를 알아듣게 하기 위해서 핵심어 검출(keyword spotting)¹⁾ 기반의 음성 인식 기술을 사용한다. 이는 사람의 음성을 계속 듣고 있다가 특정 키워드가 발성 되었는지를 검출하는 방법이다. 예를 들어 “헤이, 카카오, 신나는 노래 들어”라고 말하면, [그림 3]과 같이 연속으로 입력되는 음성 구간에서 키워드(헤이, 카카오)에 해당하는 발음 시퀀스(sequence)가 순차적으로 들어오는지를 지속적으로 감지한다.



발화자의 음성에서 핵심어 검출하는 방법

1) 발화자 음성에서 특징 추출

앞서의 과정을 조금 더 구체적으로 살펴보자. 핵심어 검출은 [그림 4]와 같은 순서로 동작한다. 우선 발화자의 음성(raw speech)에서 특징 벡터(feature)를 추출한다. 특징 벡터는 음성의 특성을 잘 반영하고 편리한 계산을 위해 사용된다. 일반적으로 사용되는 특징 벡터에는 멜 주파수 캡스트럼(Mel frequency cepstral coefficients,

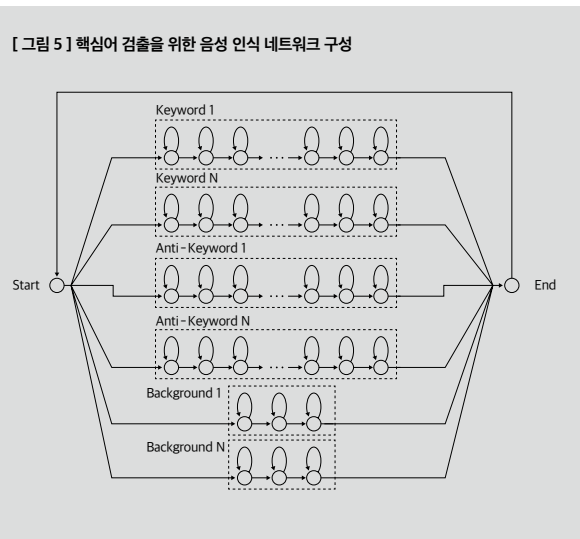
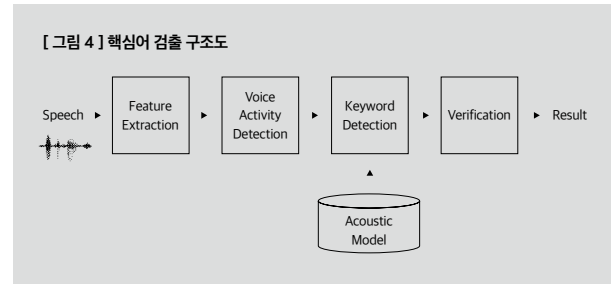
글 | 박중세 carlos.park@kakaocorp.com

어쩌다 보니 음성 관련 개발만 10년 넘게 했는데, 과거에 비해 요즘은 음성 기술이 정말 많이 대중화 되었다는 걸 실감합니다. 하지만 그래도 그 끝은 없는 분야인 것 같습니다.

글 | 정대성 denzel.jung@kakaocorp.com

10여 년 간 음성 인식 연구와 서비스 개발을 하고 있습니다. 최근 인공지능 기술이 음성 인식에 접목됨으로써 인식 품질이 개선되고 다양한 서비스에 적용되어 재미와 보람을 느끼고 있습니다. 카카오의 모든 서비스에 음성 인식 기술이 적용될 때까지 열심히 일하는게 제 목표입니다.

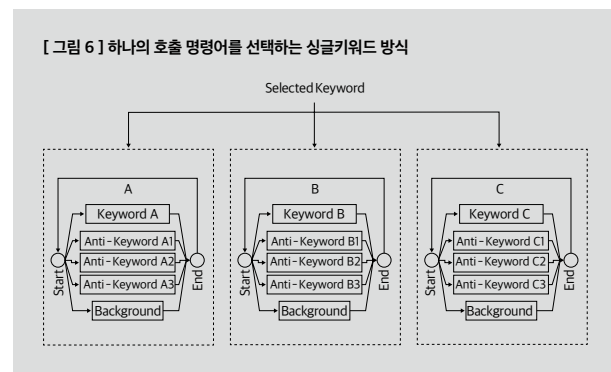
MFCC)이나 필터 बैं크 에너지(filter bank energy) 또는 지각 선형 예측(perceptual linear prediction, PLP) 등이 있다.



카카오미니의 핵심어 검출 엔진

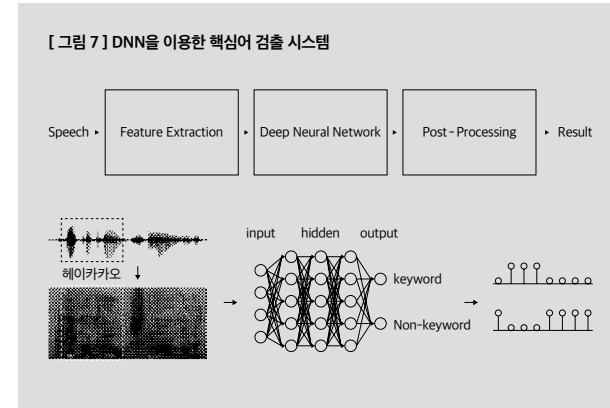
카카오미니의 핵심어 검출 엔진은 안티키워드 모델과 DNN 기법을 적용하여 높은 인식률을 가진다. 저전력, 빠른 응답 그리고 적은 메모리 사용을 위한 알고리즘 최적화 과정이 수행되었다.

대부분의 인공지능 스피커는 몇 개의 호출 명령어 중에서 하나를 선택하는 방식을 사용한다. 이 경우 실제로 동작하는 키워드는 한 개가 되며, [그림 6]처럼 싱글키워드(single keyword)를 검출하는 형태로 동작한다. 만약 사용자가 호출 명령어로 '키워드 A'를 선택하면 그에 대응되는 안티키워드 A1, 안티키워드 A2 그리고 안티키워드 A3이 함께 적용되며, 호출 명령어를 변경하면 그에 맞는 안티키워드 목록도 함께 변경된다.



앞에서 음소 기반 음성 인식 네트워크를 구성하는 방법에 대해 살펴보았는데, 음성 인식 네트워크를 구성하지 않고 더 간단하게 핵심어를 검출하는 방법도 있다. [그림 7]은 DNN을 이용한 핵심어 검출 시스템을 보여준다. 입력되는 음성에서 10msec(milliseconds) 단위의 프레임 구간마다 특징 벡터를 추출한 뒤 이를 바로 DNN 입력(input)으로 넣는다. 이때 특징 벡터는 앞뒤로 여러 프레임의

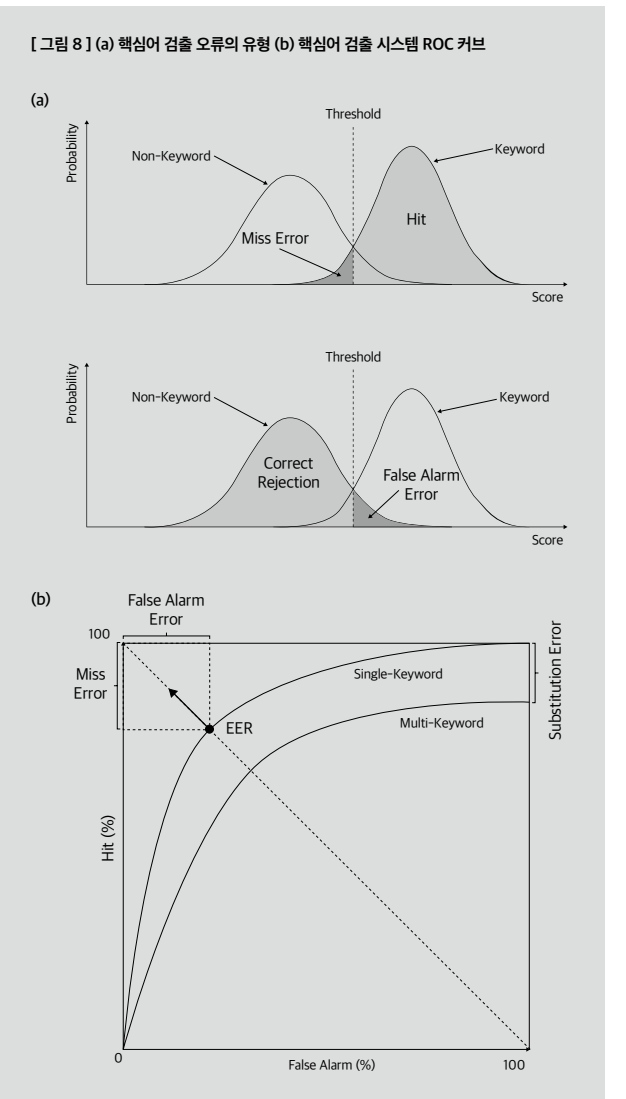
값을 합친 슈퍼 벡터(super vector) 형태로 구성한다. 그리고 입력된 특징 벡터는 미리 학습한 DNN의 계산을 통해 keyword 노드와 non-keyword 노드의 확률 값으로 변환된다. 프레임마다 발생하는 이 출력 값에 대해 후처리 과정을 거치면 키워드 검출이 가능하다. 이 방법은 백그라운드 모델을 설계할 필요가 없고 구조가 간단하다는 장점이 있다. 또한 오류가 발생한 음성 데이터를 재학습할 수도 있다. 하지만 학습 데이터를 음소 단위로 쪼개어 사용하지 않기 때문에 키워드 전체가 포함된 학습용 음성 데이터가 많이 필요하다.



인공지능 스피커의 음성 인식은 아직 완벽하지는 않지만 더욱 기술이 발전하면 지금보다 더 좋아질 것이다.

현재의 한계, 그리고 발전을 위한 노력

[그림 8] (a)는 핵심어 검출 동작 시 발생하는 오류 유형을 그림으로 표현한 것이다. 키워드를 정확히 말했다 때 핵심어 검출 프로그램의 스코어가 임계치(threshold)보다 크면 검출 성공(hit)이 되고, 스코어가 낮으면 검출 실패(miss)가 된다. 반대로 키워드가 아닌 말을 한 경우에는 스코어가 낮게 나타나 정확히 거절(rejection)이 되어야 한다. 그러나 가끔씩 키워드와 유사한 발음이 포함되면 스코어가 높게 나타나 잘못 검출되는 상황(false alarm error)이 발생한다. 이때 임계치를 크게 하면 잘못 검출될 확률(false alarm rate)은 줄어들지만 성공 확률(hit rate)도 낮아지고, 반대로 임계치를 작게 하면 성공 확률이 높아지지만 잘못 검출될 확률도 함께 커지게 된다.



[그림 8] (b)는 핵심어 검출 시스템의 특성을 살펴보기 위해 많이 사용되는 수신자 조작 특성(receiver operating characteristic, ROC) 곡선이다. 곡선 상에서 성공 확률이 증가할수록 잘못 검출될 확률도 함께 증가하는 것을 볼 수 있다. 이때 성공 확률과 잘못 검출될 확률이 같아지는 지점을 동일 오류율(equal error rate, EER)이라고 하며 시스템의 성능 지표로 많이 사용한다. 개발자의 궁극적인 목표는 ROC 곡선에서 EER이 낮아지는 방향으로 이동하여 0으로 수렴하게 하는 것이다. 그러나 실제로는 사용자의 목소리 차이, 발생 패턴 차이, 주변 잡음 등 여러 가지 성능 저하 요소가 존재한다. 카카오의 개발자들은 사용자의 음성을 더 똑똑하게 잘 알아듣는 카카오미니를 만들기 위해 더욱 정교하고 진보된 음성 인식 알고리즘을 연구하고 많은 실험 과정을 거치고 있다.

음성 인터페이스로서 카카오미니의 발전 방향들

지금보다 더 쉽고 편리하게 카카오미니를 사용하는 방법은 없을까? 카카오미니의 서비스가 다양해짐에 따라 음성 인터페이스에서의 새로운 요구 사항도 많아질 것이다. 여러 가지 발전 방향들 중에서 몇 가지를 꼽아보면 개인화, 멀티 플랫폼, 멀티 모달, 상황 인지가 있다.

개인화(personalization)

사람의 외모나 성격, 취향 등이 모두 다른 것처럼 사용자마다 목소리 특성, 사용 방법, 선호하는 호출 명령어도 다르다. 이러한 개인의 차이를 반영하는 것도 중요한 과제다. 인공지능 스피커에 입력된 사용자의 목소리를 스피커 내에서 바로 학습해서 해당 사용자의 호출 명령어를 더 잘 인식하도록 하는 방법도 많이 연구되고 있다.³

보통 카카오미니를 집에서 사용하기 때문에 본인뿐 아니라 가족들도 함께 쓰는 경우 카카오톡 메시지나 일정을 확인할 때 보안 문제가 생길 수 있다. 따라서 사용자의 목소리에만 반응하게 하는 화자인식 기술이 요구된다.⁴ 가족 중 누구의 목소리인지 구분되면 이를 통해 인증이나 추천 등의 개인별 맞춤 서비스가 가능하다. 이렇게 화자인식을 적용할 때 매번 부르는 호출 명령어를 활용하는 것이 도움이 된다. 사용자가 미리 본인 목소리로 “헤이, 카카오!”를 등록하면, 이후에도 항상 해당 호출 명령어를 말하기 때문에 해당 사용자의 목소리가 맞는지 판단하기 쉬워진다.

또한 지금은 미리 정해진 몇 개의 호출 명령어 중에서 하나를 선택하는 방식을 사용하지만, 사용자는 자신이 호출 명령어를 직접 정하고 싶어한다. 이 기능이 실제로 가능해 진다면 작명의 재미가 생기는 것이다. 어쩌면 “개똥아”, “돌머리”처럼 짓궂은 단어를 많이 입력할지도 모르겠다. 다만 사용자가 설정한 호출 명령어가 음성 인식 기술로 검출하기 어렵거나 오인식을 많이 유발하는 경우에 대한 대책도 반드시 필요하다. 해당 단어를 인식하는 데 문제가 없는지 개발자가 미리 테스트해 볼 수 없기 때문이다.

멀티 플랫폼(multi-platform)

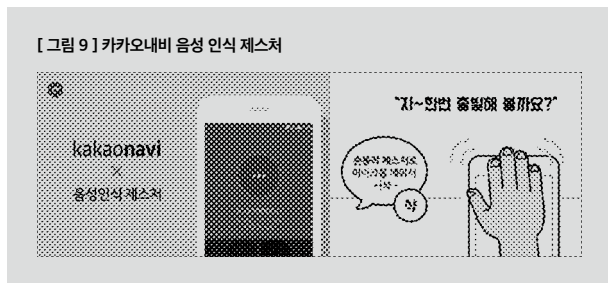
최근 인공지능 스피커에 많은 기능과 서비스가 탑재되다 보니 하나의 플랫폼에서 모든 서비스를 처리하는 게 쉽지 않다. 사용자가 환경에 따라 각각 다른 음악 서비스나 스마트홈 서비스를 이용하기 때문이다. 또한 각 업체의 인공지능 플랫폼이 잘 하는 영역도 모두 다르다. 이런 상황에서 하나의 플랫폼보다는 여러 플랫폼이 연동하는 방식도 고려할 수 있다. 애플 시리(Siri)가 수학 연산이나 질의응답 도메인에 대해서 또 다른 인공지능 플랫폼인 울프럼 알파(Wolfram Alpha)⁵를 연동하여 답변해주는 것은 오래된 사례다. 또 다른 방식으로 스마트폰이나 가전 제어는

빅스비(Bixby)를 이용하고, 카카오톡 보내기, 음악 재생, 길찾기 등은 카카오(아이)를 이용하는 경우도 상상해 볼 수 있다. 이때 사용자는 호출 명령어를 “하이, 빅스비”와 “헤이, 카카오!”를 선택해서 부를 수 있을 것이다. 또한 똑같은 명령어라도 호출 명령어에 따라 다르게 동작하도록 설정할 수도 있다.

멀티 모달(multi-modal)

음성 외에 버튼, 센서, 카메라 등 다른 입력 수단을 이용하여 스마트 기기를 웨이크업 하거나 이를 음성 인식과 함께 사용하는 방식도 많다. 손목을 위로 들면 스마트 위치에 탑재된 음성 비서가 바로 실행되거나, 집안에 있는 카메라를 똑바로 쳐다보기만 하면 사용자의 얼굴을 인식하여 “장동건 님, 무엇을 도와드릴까요?”라고 물어볼 수도 있다.

현재 카카오의 길 찾기 어플리케이션인 카카오내비에는 음성 인식 기능이 탑재돼 있어서 목적지를 음성으로 검색할 수 있다. [그림 9]처럼 근접 센서를 이용하여 스마트폰 근처에 손을 가져가면 바로 음성 인식이 동작한다. 이것도 제스처를 이용한 멀티모달의 한 가지 적용 예다.⁶



상황 인지(Context Aware)

이심전심(以心傳心)이라는 말처럼 카카오미니가 내 마음을 알아주면 참 편리할 것이다. 말하지 않아도 외출할 때면 전등을 꺼주고, 요리할 때는 타이머를 설정해주고, 잠이 들면 음악도 꺼지게 하는 것처럼 말이다. 이처럼 눈치 100단 수준의 카카오미니라면 목 아파게 “헤이, 카카오!”를 부르지 않아도 되지 않을까. 쉽지는 않겠지만 앞으로 조금씩은 눈치가 생기지 않을까 싶다. 지금의 카카오미니는 사용자 명령에 대한 답변이 끝나고 나면, 호출 명령어를 부르지 않고도 연속으로 명령할 수 있도록 설계되어 있다. 이는 사용자가 연달아서 질문하는 경우가 많다는 점을 고려한 음성 인터페이스이다.

³ 논문 | Rose, R. & Paul, D. (1990). A hidden Markov model based keyword recognition system, Proceedings of the International Conference on Acoustics, Speech and Signal Processing. , 논문 | Wilpon, J., Rabiner, L., Lee, C. and Goldman, E. (1990). Automatic recognition of keywords in unconstrained speech using hidden Markov models, IEEE Transactions on Acoustics, Speech, and Signal Processing. ⁴ 논문 | Chen, G., Parada, C. and Heigold, G. (2014) Small-footprint Keyword Spotting using Deep Neural Networks, Proceedings of the International Conference on Acoustics, Speech and Signal Processing ⁵ 논문 | Kumatani, K., et al. (2017). Direct Modeling of Raw Audio with DNNs for Wake Word Detection, Automatic Speech Recognition and Understanding Workshop ⁶ 참고 | 김명재. (2017). 카카오미니는 어떻게 목소리를 인식할까, 카카오AIR리포트 ⁷ 참고 | 스티븐 울프럼이 만든 Knowledge Engine (www.wolframalpha.com) ⁸ 참고 | 카카오 기업 블로그 blog.kakaocorp.co.kr/654

카카오미니는 어떻게 점점 더 똑똑해지는가

올 1월 말에서 2월 초에 걸쳐 카카오미니가 더욱 다양한 사람의 언어를 이해하기 시작했다. 이제 카카오미니로 치킨, 피자 등을 주문할 수 있을 정도가 됐다. ‘윤식당’이 몇시에 하는지, 현재 예매 순위 1위가 어떤 영화인지도 카카오미니에게 물어보면 알 수 있다. 현재의 실시간 이슈가 무엇인지, 주말에 손흥민이 프리미어리그 경기 라인업에 포함되었는지도 “헤이, 카카오!”만 부르면 알 수 있다. 집에서 회사까지 가장 빠르게 갈 수 있는 대중 교통정보를 알려주는 것은 물론, 카카오T를 통해 택시를 불러주기도 한다.

카카오미니의 업데이트된 서비스 명령어

- 카카오톡 주문하기
“헤이, 카카오! 후라이드 치킨 시켜줘”
“헤이, 카카오! 오늘 할인 메뉴 뭐야?”

- 어학/사전
“헤이, 카카오! 냉장고가 영어로 뭐야?”
“헤이, 카카오! 점입가경이 무슨 뜻이야?”

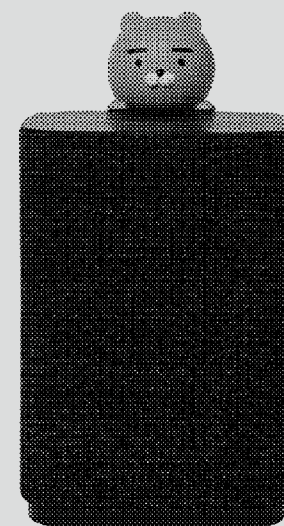
- 카카오 T 택시
“헤이, 카카오! 강남역 가는 택시 불러줘”
“헤이, 카카오! 택시 언제 와?”

- 프리미어리그 경기 일정/결과
“헤이, 카카오! 맨유 경기 일정 알려줘”
“헤이, 카카오! 어제 손흥민 잘했어?”

- TV/영화
“헤이, 카카오! 윤식당 언제 해?”
“헤이, 카카오! 요즘 재밌는 드라마 뭐있어?”
“헤이, 카카오! 1987 관객수 얼마야?”

- 교통정보
“헤이, 카카오! 9007번 버스 언제와?”
“헤이, 카카오! 강남역까지 얼마나 걸려?”
“헤이, 카카오! 이태원에서 명동까지 얼마나 걸려?”

- 실시간 이슈/정보
“헤이, 카카오! 실시간 이슈 알려줘”
“헤이, 카카오! 캐나다 수도 어디야?”
“헤이, 카카오! 라면 칼로리 얼마야?”



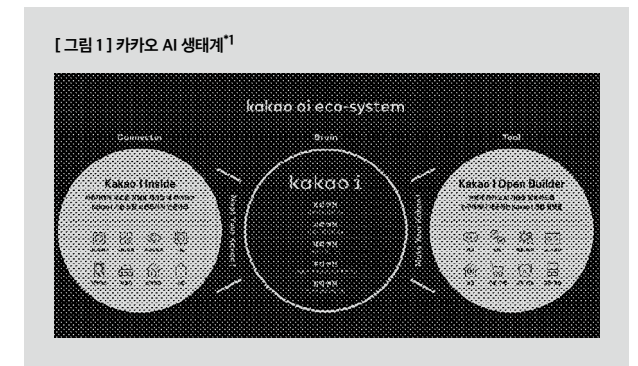
카카오미니는 점점 똑똑해지는 중이다

상상력 가득한 공상과학(SF) 영화를 보며 기대했던 AI의 모습은 아닐지라도 카카오미니는 점점 똑똑해지고 있는 중이다. 카카오미니를 사용하고 있는 사람이라면 아마도 한 번쯤은 카카오미니가 이전에는 알아듣지 못했던 명령을 어느 순간 이해하고 대답하는 현상을 경험했을 것이다. 그렇다면 카카오미니는 어떻게 진화하는 것일까?

<카카오 시리포트>를 통해 많이 소개된 딥러닝(deep learning) 기술은 카카오미니가 계속해서 학습하고 진화하도록 도와준다. 기술적인 내용은 이전의 <카카오 시리포트>를 통해 많이 소개돼 왔기에, 이번 글에서는 서비스 기획자의 시각으로 카카오미니가 사람의 언어를 이해하는 과정을 소개하고자 한다.

카카오미니는 어떻게 사람의 언어를 이해하는 것일까

카카오미니의 두뇌 역할을 하는 카카오(아이)의 코어 엔진(core engine)은 스피커 안에 있는 것이 아니다.



카카오미니에는 듣는 역할을 하는 ‘마이크’와 사용자에게 말을 하는 역할을 ‘스피커’가 탑재되어 있고, 사용자가 말한 명령은 인터넷망을 통해 카카오 코어 엔진 서버로 전송된다. 코어 엔진은 사용자가 인지할 수 있는 ‘업데이트’ 과정 없이 지속적으로 개선되기 때문에, 마치 사람처럼 학습한다고 느껴질 수 있다.

사용자가 “헤이, 카카오!”라고 부르면 카카오미니는 땡 소리와 함께 노란색 LED를 점등하며 대화 가능한 상태가 된다. 그 때부터 입력되는 음성은 서버로 전송되고 음성 엔진은 이를 텍스트로 변환한다. 대화 엔진은 변환된 텍스트를 사용자가 어떤 ‘의도’로 말을 했는지 ‘분류’하고, 각각의 의미 있는 단어를 ‘추출’한다. 서비스 에이전트는 추출된 단어를 이용하여 사용자 ‘의도’에 맞는 적절한 답변을 만들고, 이 답변이 다시 스피커로 전송되어 출력된다.

카카오미니가 언어를 잘 이해하려면 전송된 ‘음성을 텍스트로 변환’하는 과정과 텍스트를 의도에 맞게 ‘분류’하고 의미 있는 단어를 ‘추출’하는 것이 중요한데, 이때 ‘딥러닝’이라는 기술이 사용된다.

글 | 김다현 lea.kim@kakaocorp.com

쇼핑, 사전, 동영상과 같은 버티컬 서비스의 검색 기획을 해왔습니다. 2017년 초 AI 서비스팀에 합류한 후 카카오미니의 탄생을 함께 했습니다. 특히 카카오(아이)가 좀 더 똑똑해져서 많은 사람들에게 더욱 필요한 서비스가 되었으면 합니다.

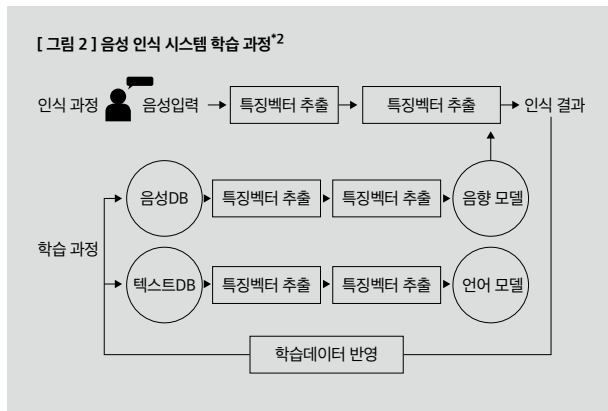
카카오미니가 학습하는 방법

카카오미니를 많이 사용할수록 음성 인식 성능은 점점 더 좋아진다.

그 원리는 무엇일까? 아래 그림은 카카오 음성 엔진의 학습

방법이다. 음성 인식 시스템은 크게 ‘인식’ 과정과 ‘학습’ 과정으로

나눌 수 있다.



‘인식 과정’은 사용자의 음성을 프로그램이 인식하여 텍스트 결과를 내어주는 것으로, 실제로 사용자의 눈 앞에서 이루어지는 과정이다.

이때 음성 엔진은 미리 학습된 음향 모델과 언어 모델을 이용한다.

‘학습 과정’은 사용자의 음성과 인식된 텍스트를 다시 학습하여 음향 모델과 언어 모델을 만드는 단계다. 실제 사용자 데이터는 다양한 연령, 목소리 특성, 주변 환경을 포함한 데이터이므로 이를 학습하여 새로운 모델을 만드는 것은 성능 개선에 필수적인 요소이다. 일반적으로 음향 모델은 깊은 신경망(deep neural network, DNN) 방식을 적용하면 대용량의 사용자 데이터를 이용하여 더욱 정교한 모델링이 가능하다.

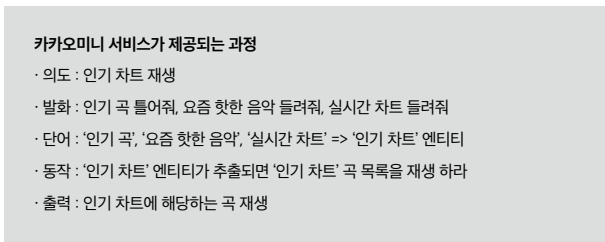
‘의도’ 분류를 학습하기 위해서는 ‘충돌’하는 명령어가 없는지 검토하는 데이터 전처리 과정이 중요하다. 유입된 명령어를 ‘의도’에 따라 분류한 후 ‘공통’으로 등장하는 것들은 ‘공통 명령어’로 묶어 사용자의 이전 상태에 따라 적절한 ‘의도’로 분류한다.

예를 들어 ‘그만’이라는 명령어가 유입되었을 때를 생각해 보자. 이 경우 사용자가 이전부터 ‘음악’을 듣고 있었다면 ‘그만’은 ‘음악 종료’에 해당하는 ‘의도’이고, ‘라디오’를 듣고 있었다면 ‘라디오 종료’에 해당하는 ‘의도’로 분류하는 방식이다. 정제된 데이터는 말뭉치에 등장하지 않은 단어 중 말뭉치에 존재하는 것과 유사한 단어를 학습하기 위해 단어 임베딩(word embedding) 방식을 적용하여 모델링 한다. 최근에는 GloVe(Global Vectors for Word Representation)를 활용하여 학습하고 있다. 특히 사용자들이 카카오미니와 더 많은 대화를 할수록 카카오미니는 점점 더 똑똑해진다.

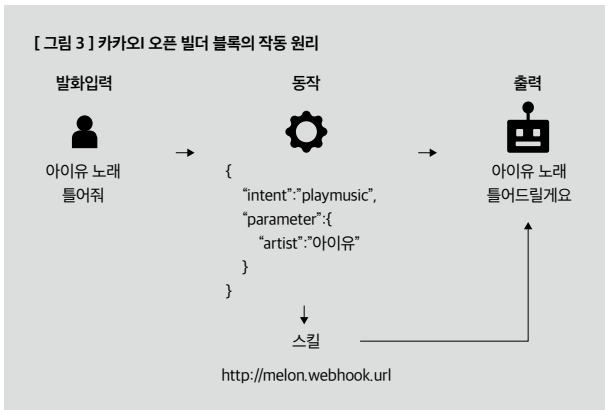
카카오미니 서비스가 추가되는 과정

카카오미니 서비스가 만들어지는 과정은 기존의 웹서비스와는 조금 다르다. 어떤 서비스를 만들지 결정하면 서비스를 사용하는 ‘의도’부터 분류한다. 예를 들어 ‘음악 서비스’를 만든다고 하자. 음악을 듣는 형태는 다양하다. 인기 차트를 듣고 싶을 수도 있고, 분위기에 맞는 음악이나 특정 가수의 곡을 듣고 싶은 의도도 있다. 음악을 듣다가 마음에 안들면 다음 곡 재생을 요청할 것이고 이것 저것 요청하는게 귀찮다면, 그냥 ‘음악 좀 틀어’라고 말할 수도 있다.

사용자의 ‘의도’를 분류한 이후에는 그 의도를 이루기 위해 사용자가 명령할 만한 ‘발화’를 만든다. ‘발화’는 의미 있는 ‘단어’의 조합으로 구성되고 특정 단어의 조합에 따라 서비스 에이전트가 어떤 동작을 수행할지 정의한다. ‘동작’ 수행은 카카오미니의 캐릭터에 맞게 출력 문장을 다듬는 과정을 거치고, 최종 결과물이 스피커를 통해 ‘출력’된다.



이 모든 과정은 카카오의 개발 플랫폼인 ‘카카오 오픈 빌더(Kakao I Open Builder)³’를 통해 이루어진다.

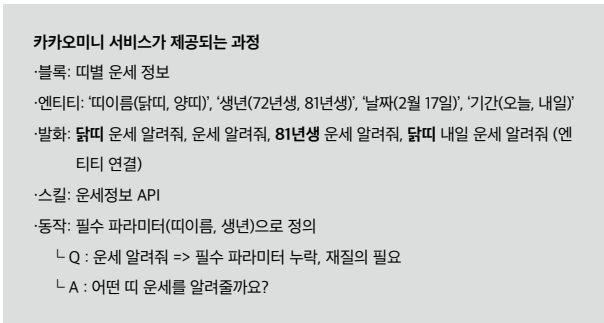


카카오 오픈 빌더는 ‘엔티티(entity)’, ‘블록(block)’, ‘스킬(skill)’로 구성되어 있다. ‘엔티티’는 사용자의 의도에 맞는 동작을 수행하기 위해, 주요 단어를 ‘추출’하도록 규정한 데이터 사전이다. ‘블록’은 사용자 명령으로 부터 의도를 파악하여 서비스 에이전트가 수행할 액션과 응답할 내용을 형식에 맞게 설계하고 정의하는 모듈(module)이다. ‘스킬’은 ‘블록’에서 사용자의 ‘의도’를 수행하기 위해 다양한 API(Application Programming Interface)를 호출하여 동작을 수행할 수 있도록 한다.

기획자는 위에서 정의한 ‘의도’를 빌더(builder)의 ‘블록’으로

생성하고 사용자 발화에 포함되는 ‘단어’를 ‘엔티티’ 사전으로 등록한다. ‘블록’에서 ‘발화’를 입력하면, 빌더는 입력된 발화의 단어를 사전에 등록된 ‘엔티티’와 연결한다.

‘블록’에서는 설정한 ‘스킬’을 연동할 수 있도록 ‘동작’ 파라미터(parameter)도 정의한다. ‘엔티티’가 제대로 추출되었는지 검증하고, 필수 ‘엔티티’를 정의해서 필수 값이 비어 있다면 사용자에게 ‘재질의’하도록 설정할 수 있다.



카카오미니가 매일 조금씩 좋아지는 이유

카카오는 자주 유입되는 발화와 의도를 분석할 수 있는 통계 시스템을 가지고 있다. 서비스 기획자는 이를 통해 매일 유입되는 데이터를 모니터링할 수 있다. 자주 들어오는 명령은 어떻게 더 스마트하게 답변할 수 있을지 검토하고, 실패한 명령은 왜 실패를 했는지 그 원인을 파악한다.

실패의 원인은 앞에서 설명한 ‘음성을 텍스트로 변환’ 하는 과정일 수도 있고, 텍스트를 ‘분류’하거나 ‘추출’하는 것이 잘못 되었을 수도 있다. 실패의 원인이 파악되면 이는 카카오 코어 엔진을 업데이트하거나 카카오 오픈 빌더에서 엔티티, 블록, 스킬을 수정해서 해결한다.

수정된 내용은 지속적으로 서버에 업데이트되기 때문에 카카오미니가 사용자의 의도를 한번에 이해하지 못했다고 해서 너무 실망하지 않아도 된다. 차츰 시간이 지나면서 내가 알지 못하는 사이에 끊임없이 업데이트될 것이기 때문이다.

앞으로 카카오는 어떻게 진화하게 될까

카카오는 카카오미니 외에 다양한 기기에 탑재 될 수 있다.

카카오미니가 아닌 다른 형태의 스피커나 냉장고, 에어컨과 같은 가전 제품이 될 수도 있다. 자동차나 집안의 스마트홈 시스템에도 카카오가 들어갈 수 있다. 그렇게 되면 운전하면서 “헤이, 카카오!”를 불러 길을 찾거나 음악을 들을 수 있고 음성 명령만으로도 집안의 불을 켜거나 끌 수도 있을 것이다.

또한 다양한 서비스를 추가할 계획도 가지고 있다. 카카오 통계 시스템 분석을 통해서 자주 유입되는 명령어 중 ‘서비스’가 없어서 대응하지 못하는 기능부터 업데이트 할 계획이다. 카카오 번역⁴, 보이스톡, 장소 정보 제공 등의 서비스를 준비 중이고, 외부 제휴를 통해 다양한 콘텐츠 서비스 또한 제공할 예정이다.

카카오 시스템은 기존 카카오미니 사용자들에게도 동일한 업데이트를 지원할 예정이다. 향후 카카오가 탑재되는 자동차, 스마트홈, 스피커 등은 비록 기기의 형태와 기능은 다르지만 사용자들에게 모두 동일한 경험을 제공 할 것이다.

¹ 참고 | <https://kakao.ai/about> ² 참고 | 김훈, “음성 인식 방법과 카카오의 음성형 엔진,” 카카오시리포트 Vol.6, 2017 ³ 참고 | <https://i.kakao.com/openbuilder/#/login> ⁴ 참고 | <https://translate.kakao.com>

카카오브레인의 1년 간의 발자취



Kakao Brain	이수경 인치원 카카오브레인의 1년, 그리고 미래	44
section	이수경 안다비 음성은 미래의 인터페이스다	50
	이수경 고독의 시대와 AI	54

카카오브레인의 연구 기술을 설명하는 Brain's Pick, 대중문화 속 AI를 살펴보는 AI in pop-culture 외, 이번 호에서는 카카오브레인의 어제와 오늘, 그리고 내일에 대해 설명한 글을 준비하였습니다. 이 글 안에는 카카오브레인이 걸어온 길과 현재 진행 중인 주요 프로젝트, '자율·공개·협력'이라는 카카오브레인의 3대 연구 원칙, 그리고 카카오브레인의 향후 방향이 상세히 설명돼 있습니다.

카카오브레인의 1년, 그리고 미래

최근 부상하고 있는 자율주행 기술부터 질병 진단, 재난 구조 로봇에 이르기까지 인공지능(AI)이 많은 분야에서 핵심적인 역할을 담당할 것으로 기대를 모으고 있다. 전 세계 큰 기업들이 너도나도 AI 기술과 인재에 대한 투자에 뛰어드는 이유다. 카카오 또한 지난 해 200억 원을 출자해 카카오브레인이라는 인공지능 연구소를 자회사로 세웠다. 오랜 시간 연구가 필요한 분야인 만큼, 카카오브레인은 지난 1년 간 원활한 연구 환경과 문화를 조성하는 데 힘써 왔다.

kakaobrain

개요	
대표이사	김범수
설립일	2017.02.01
소재지	경기도 성남시
직원 수	30명
연혁	
2017.02	카카오브레인 법인 설립
2017.04	초지능연구센터(Center for Super Intelligence) 산학협력 협약
2017.05	인공지능 기반 개인화 플랫폼 기업 ‘스컬터랩스’ 투자 (케이큐벤처스 공동 투자) 카카오브레인-한국기원 딥러닝 오픈리서치 MOU 체결
2017.06	로봇 모듈 플랫폼 기업 ‘럭스로보’ 투자(카카오 인베스트먼트 공동 투자)
2017.07	제주 텐서플로(Tensorflow) 머신러닝 캠프 공동 주최
2017.08	로봇 스타트업 ‘토록’ 투자(케이큐벤처스 공동 투자)
미션	
인간처럼 생각하고 행동하는 기계(machine)를 통해 인류가 이제까지 풀지 못했던 난제에 도전합니다.	

글 | 이수경 samantha.lee@kakaobrain.com

2016년 3월 알파고와 이세돌 9단이 펼치는 세기의 대결을 두 눈으로 목도한 이후 인공지능을 제대로 공부해봐야겠다고 결심했습니다. 인공지능 본진이자 연구소인 카카오브레인으로 걸어 들어온 이유입니다. 인공지능 기술과 이로 인해 바뀔 미래 사회를 다루는 글을 통해 사람들과 소통하고 싶습니다.

글 | 인천원 ryan.in@kakaobrain.com

돌아온 탕아. 학부생 때 인공지능(AI)을 전공했고 이런저런 이유로 사바 세계를 헤맸습니다. 결국, AI 사업 전략을 세우는 ‘전략 쟁이’로 활약하고자 카카오브레인에 합류했습니다. 글로벌 AI 동향을 보며 이따금 멘붕에 빠지기도 하지만, 대한민국 AI 대표 주자가 되겠다는 책임감으로 제 앞에 놓인 역경을 이겨내고 있습니다. 매일 내딛는 걸음 걸음이 언젠가 다리를 정상으로 가는 길이기를 소망합니다.

카카오브레인의 연구 개발 문화

[그림 1] 카카오브레인의 연구 개발 방식 중 하나인 버스 시스템을 시각화



버스 시스템은 연구원들이 자율적으로 마일스톤(프로젝트 일정 관리 이징표)을 세우고 단기간에 집중적인 성과를 내고자 만든 카카오브레인만의 연구 개발 방식이다. ‘좋은 사람들과 함께하는 긴 여정을 위한, 조금 더 즐거운 여행길’이라는 컨셉을 내세운다. 프로젝트 리더인 ‘드라이버’가 새로운 연구 주제를 발제하면, 여기 흥미를 느낀 연구자들이 ‘승객’으로 모여 4주간 자율적으로 연구한다. 외부와의 협업을 진행할 수도 있다. 여기서 핵심은 보고가 아닌 ‘공유’다. 모든 구성원이 각자 어떤 연구를 하는지, 왜 하는지, 어느 단계까지 왔는지 안다면 적시에 필요한 자원을 제공할 수 있다.

미니 컨퍼런스에서는 지난 기간 진행된 연구에 대한 발표가 이뤄진다. 주요 프로그램은 크게 세 가지로 구분된다. 모든 프로젝트 내용을 5분으로 요약하는 ‘스포트라이트’ 세션, 프로젝트의 목표와 해결 과제, 활용된 기술 등에 관해 상세히 소개하는 ‘구두 발표’ 세션, 프로젝트 주요 내용을 포스터 형태로 인쇄해 놓은 뒤 관객과 1:1 토크를 나누는 ‘포스터’ 세션이 바로 그것이다. 여러 번의 시행 착오 끝에 현재는 프로젝트 리더인 드라이버, 프로젝트 참가자인 승객을 중심으로 자율 발제 시스템이 안정적으로 운영되고 있다.

카카오브레인은 연구원들이 최신 동향을 파악하고 더 나은 기술을 체득하는데 필요한 학습 비용 일체를 지원한다. 국내외 학회 참가비나 연구 장비, 도서 등이 여기에 속한다. 물론 출판 전 논문을

공개하는 사이트인 아카이브(arXiv) 덕분에 학회가 열리기 전에 유명한 논문을 먼저 읽어 볼 수 있고, 발표 동영상까지 볼 수 있는 시대다. 그렇기에 굳이 큰 비용과 귀중한 시간을 들여서 학회에 가야만 하느냐는 의문이 들 수 있다. 그러나 ‘백문이 불여일견(百聞不如一見)’이라는 속담에서 보듯이, 현장에서 직접 보고 느끼며 지적인 자극을 받는 것은 단순히 논문을 읽고 이해하는 것과는 비교할 수 없는 가치가 있다. 전 세계에서 인공지능을 연구하는 사람들과 교류하며 최신 지견을 교환할 수 있다는 것 또한 장점이다.”

카카오브레인의 3대 원칙 : 자율, 공개, 협력

[그림 2] 카카오브레인 연구원들



카카오브레인이 자체적으로 내세우는 3가지 원칙이 있다.

첫 번째는 '자율'적으로 연구한다. 기업에 소속된 연구 조직은 통상적으로 일관된 목표를 가지고 한 가지 기초 연구에 오랜 시간을 투자하기가 쉽지 않다. 또는 기업 경영진이 만드는 회사의 전략에 따라 연구 조직이 쉽게 와해하거나 격변을 겪을 수도 있다. 따라서 인공지능 분야에서 유의미한 성과를 내려면 장기 연구를 보장해야 한다. 아울러 연구자를 중심으로 연구·개발(R&D)이 가능한 환경적 토대를 조성해야 한다. 카카오브레인에서는 프로젝트 주제를 자유롭게 선정하고, 연구에 필요한 모든 것을 지원하는 것을 원칙으로 삼는다. 책임자가 권한을 갖고 프로젝트에 관한 의사 결정을 주도하는 시스템이다.

두 번째는 기술을 '공개'한다. AI를 주도하는 많은 글로벌 기업은 공개 연구를 지향하고 있다. 인재 풀이 적기도 하거니와, 극소수 선구자들이 학교에 적을 두고 있어서 나타난 현상이기도 하다. 최근에는 학교보다도 더 많은 논문이 기업에서 나오기까지 한다. 기업이 학교보다 더 많은 데이터와 더 강력한 컴퓨팅 파워를 갖출 수 있었던 덕분이다. 이처럼 기업들이 인공지능 분야 공개 연구를 주도하면 혁신의 시기를 앞당길 수 있다. 카카오브레인은 자체적으로 진행한 연구 내용을 공개하고, 관련 기술을 오픈 소스화하는 방향을 추구하고 있다.²

세 번째는 커뮤니티와 '협력'한다. 신선한 재료는 좋은 요리의 가장 중요한 조건이다. 제 아무리 세계 최고의 실력을 갖춘 셰프라도 나쁜 재료로 맛있고 몸에 좋은 요리를 만들기가 쉽지 않다. 마찬가지로 좋은 기술과 연구를 위해서는 기초 연구가 탄탄해야 기술 혁신을 이룰 수 있다. 진일보한 기술력을 갖추려면 혁신 주체들 간의 강한 연대, 활발한 교류를 통해 개방형 생태계를 구축해야 한다. 기초 연구를 하는 학계 전문가 파트너들과의 협력이 더욱 중요한 이유다. 이에 카카오브레인은 국내외 대기업, 중소기업, 스타트업과 긴밀하게 관계를 맺음으로써 국내 딥러닝 기술 개발과 그 인프라 구축에 앞장서고자 한다.

연구 분야·프로젝트

카카오브레인이 연구하는 분야는 머신러닝 방법론, 로보틱스 및 자율 주행, 강화 학습, 최적화, 자연어 처리(natural language processing, NLP), 음성 인식 및 합성, 의료 진단 등이 있다. 연구 주제는 특정 도메인에 국한되지 않으며, 반드시 한국에서 상용화 가능한 기술이어야 한다는 제약 조건도 없다. 또한, 카카오로부터 데이터를 받거나 카카오브레인이 개발한 기술을 카카오 서비스에 적용하는 등 카카오브레인과 카카오의 협업은 언제든지 다양한 형태와 방식으로 진행 될 수 있다.

다음은 카카오브레인 내부에서 진행되는 프로젝트 중 일부다.

브레인 클라우드(Brain Cloud)

브레인 클라우드는 아마존의 아마존웹서비스(AWS), 마이크로소프트(Microsoft)의 애저(Azure)와 같이 응용 프로그램에서 요구하는 자원과 환경에 따라 사용할 수 있는 클라우드 플랫폼이다. 현재는 카카오브레인 내부 개별 프로젝트가 요구하는 클라우드 플랫폼 기능을 제공하는 한편, 연구 속도를 가속화하고 카카오브레인의 그래픽 처리 장치(graphics processing unit, GPU) 장비의 사용률을 극대화하는 기술을 개발하고 있다.

rethink-Go

기존의 게임 방법론을 개선해 바둑을 잘 두는 강한 인공지능 플레이어를 만드는 것을 1차 목표로 삼고 있다. 사람처럼 게임을 두고 학습하는 게임 인공지능의 일반적인 요구를 만족하는 것이 그 다음 단계다. 게임뿐만 아니라 다른 문제를 풀 수 있도록 알고리즘을 확장하는 것이 최종 목표다.

WordRep

카카오브레인의 WordRep 프로젝트는 단어 벡터 사전을 이용해 단어의 유사성을 시각화하고자 진행됐다. 현재는 Wordweb³이라는 사이트를 통해 서비스 중이다. 향후 형태소 분리가 반영된 벡터 사전으로 교체하거나 후처리를 통해 동일한 단어들을 제외하는 등의 방법으로 더 나은 결과를 내기 위한 연구를 진행할 계획이다.⁴

음성 스타일 변환(Voice style transfer)

카카오브레인 연구팀은 특정인의 목소리를 제삼자의 것으로 손쉽게 변환할 수 있다면, 이 음성 변환 기술이 활용 가능한 애플리케이션의 잠재적 가치가 클 것으로 보고 관련 연구를 진행하고 있다. 신규 아키텍처를 기반으로 상용화가 가능한, 높은 수준의 음성변환 시스템 구축이 최종 목표다. 자가 학습을 통해 음성에서 특징을 추출할 수 있도록 오랜 시간과 노력을 투자한다는 계획이다.

CSI(초지능연구센터)

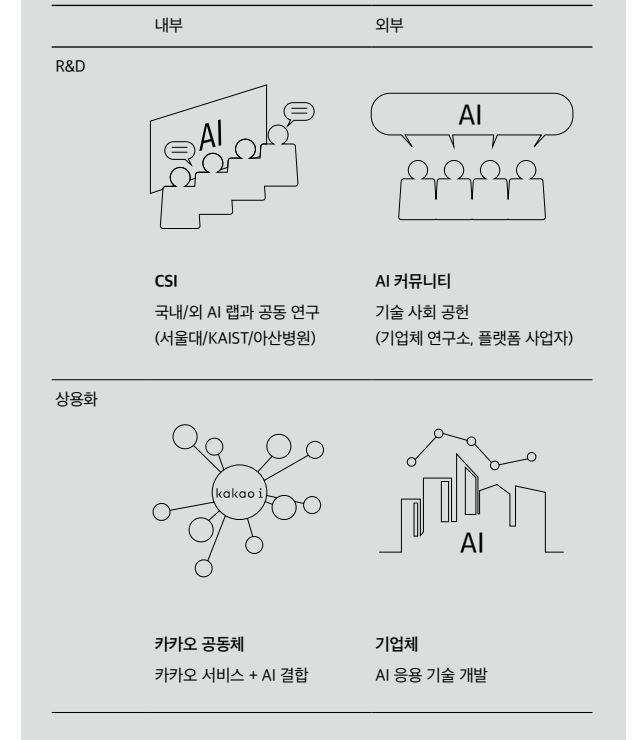
[그림 3] 카카오브레인의 초지능연구센터 설립을 위한 산학협력 체결



카카오브레인은 지난해 4월 우리 사회의 미래를 좌우할 AI기술 전문가들의 기초 연구를 집중적으로 지원하고자 CSI(Center for Super Intelligence, 초지능연구센터)와 산학협력 협약을 맺었다. CSI와 카카오브레인은 그 이후 연구 성과를 거두고 활발하게 토론하는 장을 마련하자는 취지로 연중 네 번씩 워크숍을 열고 있다.

빠른 속도로 발전하는 인공지능 분야는 개방, 공유, 협업에 기초한 집단 연구가 필요하다. 기술이 아직 성숙하지 않았고 초기 연구 단계인 만큼 산학협력을 통해 카카오의 인프라와 빅데이터 자원을 활용하는 개방형 연구(오픈 리서치) 장점을 살리면 인공지능 기술 발전을 선도할 수 있을 것으로 기대된다.

[그림 4] 카카오브레인을 주축으로 한 연구 협력 생태계



CSO가 말하는 카카오브레인

카카오브레인은 김범수 최고경영책임자(CEO)와 김남주 최고기술책임자(CTO)를 주축으로 하는 인공지능 회사로 지난해 2월 설립됐습니다. 카카오브레인의 비전에 공감한 연구원들이 한 명씩 합류해 올해 들어 30명 규모로 커진 것입니다. 작으면 작다고 할 수 있고, 크면 크다고 할 수 있는 숫자입니다. 분명한 건 인공지능 연구를 통해 가시적인 성과를 내려면 더 많은 인재가 필요하다는 점입니다. 아직도 설립 초기 단계에 있는 회사라는 의미겠지요. 따라서 충분한 역량을 갖춘, 같은 목표를 향해 달려갈 연구원이 있다면 마음껏 연구할 수 있는 환경을 지속해서 제공하는 것이 카카오브레인이 나아갈 방향이기도 합니다.

카카오브레인의 연구는 자율, 공개, 협력을 원칙으로 합니다. 다시 한번 정리하자면, '자율'은 연구원들이 자유롭게 주제를 선정하되, 스스로가 세운 원리 원칙을 토대로 연구함을 의미합니다. '공개'는 연구 과정과 성과를 카카오브레인 내외로 공유함을 뜻합니다. 마지막으로 '협력'은 카카오뿐만 아니라 관련 학계 및 업계와의 협업을 극대화함으로써 대한민국 인공지능 생태계에 기여함을 가리킵니다.

카카오브레인은 특정 목표를 달성하는 인공지능 서비스를 개발하는 것뿐만 아니라, 인간처럼 생각하는 시스템을 만드는 데 방점을 두고 있습니다. 아직 현재의 인공지능 기술 수준으로는 인간처럼 생각하는 기계(artificial general intelligence, AGI)를 구현하기란 어렵습니다. 장기적인 투자와 연구가 필요한 이유입니다. 이런 치열한 경쟁에서 우위를 점하기 위해서는 서비스의 기반이 될 비전(vision), 언어, 판단 및 추론에 필요한 기초 연구가 선행되어야 할 것입니다. 아울러 이러한 연구를 뒷받침할 클라우드 플랫폼 구축도 병행되어야 하겠지요. 특히, 국내의 우수한 연구 개발 인력들이 더 좋은 연구 환경을 쫓아 해외로 유출되는 상황을 개선하는 것은 카카오브레인뿐만 아니라 국내 모든 기업의 과제라고 생각합니다.

인공지능 분야는 이미 GAFA⁵와 BAT⁶ 등 해외 거대 플랫폼 사와 벌이는 경쟁의 각축장이 돼 버렸습니다. 연구 자본 규모로 보면 수백 배 이상 차이가 납니다. 중국에서는 정부 차원의 지원 또한 아끼지 않고 있습니다. 테슬라의 전기자율주행차, 구글 유튜브의 포털, 아마존의 아마존웹서비스(AWS) 등의 인공지능 서비스들이 기술 및 규모 격차를 이용해 국내 시장을 잠식해 나갈 가능성 또한 큼니다.

이런 상황에서 후발 주자로 나선 카카오브레인이 단독으로 경쟁을 선도하기란 매우 어렵습니다. 아울러 연구 인력 규모도 작습니다. 하지만 카카오브레인은 앞서 내세운 자율, 공개, 협력이라는 원칙을 지키면서 인공지능 연구를 통해 기술 혁신을 이루려는 연구자와 기업, 그리고 유관 부처와의 적극적인 협력 추진을 통해 앞으로 나아가고자 합니다.

카카오브레인 최고전략책임자(CSO) 인치원

*1 참고 | CVPR 2017 참관기 <https://brunch.co.kr/@kakao-it/143> *2 참고 | 세상을 바꾸고 싶다면, 딥러닝(김남주) <https://brunch.co.kr/@kakao-it/59> *3 참고 | <http://ling.kakaobrain.com/wordweb> *4 참고 | Brain's Pick : 단어 간 유사도 파악 방법 <https://brunch.co.kr/@kakao-it/189> *5 참고 | 미국 기업인 구글(Google), 애플(Apple), 페이스북(Facebook), 아마존(Amazon)을 가리킨다. *6 참고 | 중국 기업인 바이두(Baidu), 알리바바(Alibaba), 텐센트(Tencent)를 가리킨다.

음성은 미래의 인터페이스다

1) 지난해 종결한 KBS2 드라마 ‘고백부부’에는 오랜 기간에 걸쳐 쌓인 오해와 갈등으로 이혼을 결심한 진주(배우 장나라)와 반도(배우 손호준) 커플이 등장한다. 어느 날 갑자기 1999년 스무 살 대학 시절로 시간 여행을 떠난 두 사람은 젊은 날 자신들의 모습을 되돌아보며 서로의 소중함을 깨닫는다. 극 중에서 갑작스럽게 과거로 돌아간 진주는 신장염으로 투석 치료를 받다 돌아가신, 꿈에서도 그리던 엄마 은숙(배우 김미경)과 재회한다. 현재로 돌아와서도 엄마를 기억하고 싶었던 진주는 노래방에서 노래하는 엄마의 목소리를 녹음한 테이프를 책상 서랍 안에 넣어 고이 간직한다.

2) 세계 최고의 성악가로 칭송받던 루치아노 파바로티(Luciano Pavarotti). 아쉽게도 지난 2007년 췌장암으로 세상을 떠났으나¹⁾, 팬들은 콘서트 영상이나 음반 CD를 통해 11년째 그를 추억하고 있다. 전 세계적으로 큰 사랑을 받았던 파바로티도 피해 가지 못한 질타가 있었다. 바로 자신의 모국어인 이탈리아어로 쓰여진 작품으로만 활동했다는 점²⁾이다. 이탈리아어, 프랑스어, 독일어, 영어로 작성된 오페라 작품이 주를 이루는 만큼 다국어에 능통하지 못한 성악가는 치명적인 핸디캡을 가질 수 밖에 없다. 파바로티의 경쟁자였던 플라시도 도밍고(Placido Domingo)가 모국어인 스페인어는 물론 프랑스어, 이탈리아어, 독일어 등 세계 각국의 언어로 쓰여진 작품 속 배역을 맡았던 것과는 대조적인 모습이다.³⁾

만약 진주가 엄마의 목소리로 부른 노래를 수십, 수백 곡 가지고 있었다면 엄마를 향한 그리움을 달래는 데 큰 도움이 됐을 것이다. 파바로티도 마찬가지다. 이탈리아어로 작곡된 작품에 한정해 음악 활동을 펼치던 파바로티가 세계 각국 오페라 작품 속 배역을 맡았다면 활동 스펙트럼을 더 넓힐 수 있었을 것이다. 이런 발칙한 상상을 해볼 수 있는 배경에는 바로 인공지능을 활용한 음성 변환(voice conversion) 기술이 있다.

글 | 이수경 samantha.lee@kakaobrain.com

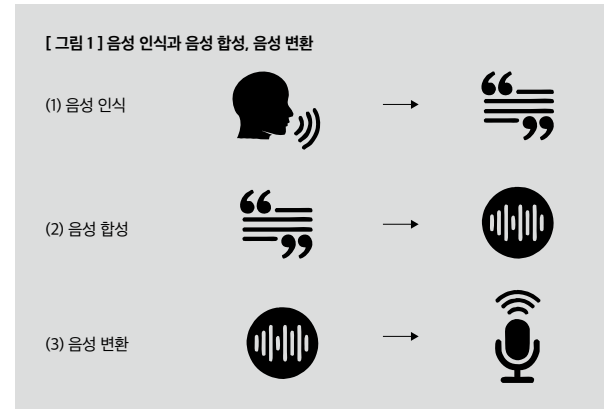
2016년 3월 알파고와 이세돌 9단이 펼치는 세기의 대결을 두 눈으로 목도한 이후 인공지능을 제대로 공부해봐야겠다고 결심했습니다. 인공지능 본진이자 연구소인 카카오브레인으로 걸어 들어온 이유입니다. 인공지능 기술과 이로 인해 바뀔 미래 사회를 다루는 글을 통해 사람들과 소통하고 싶습니다.

글/감수 | 안다비 avin.hero@kakaobrain.com

언뜻 보면 예쁜 한글 이름 같지만, 많고 크다(多叻)는 뜻을 가지고 있습니다. ‘욕심 많은’ 이름 그대로 매사에 의욕이 넘치지만 가끔은 생각에 압도 당하기도 합니다. 컴퓨터 과학을 6년 간 공부하고 엔지니어로서 5년을 살았습니다. 또 어렸을 적부터 인공지능과 같은 최첨단 기술을 좋아했습니다. 전 세계를 놀라게 한 알파고가 인공지능을 핫키워드로 만드는 것을 보며 시대를 잘 타고 났다고 느끼는 중입니다. 구글 검색, 넷플릭스처럼 지구 상의 많은 이들이 좋아할 유명 서비스를 만들고 싶은 자칭 야망가. 적당한 때를 기다리며 엔지니어로서 계속해서 열정적으로 일하고 있습니다.

음성 변환이 유력한 4개 분야

음성 변환을 이해하려면 음성 인식(speech recognition)과 음성 합성(voice synthesis)에 대한 이해가 전제되어야 한다. [그림 1]은 음성 인식과 음성 합성, 음성 변환의 차이를 보여준다. 음성 인식은 특징인이 말한 내용을 텍스트 형태로 이해하는 것이며, 음성 합성은 텍스트를 음성으로 변환함을 의미한다. 음성 합성과 음성 변환 모두 음성을 결과물로 생성한다. 그러나 음성 변환은 음성, 음성 합성은 텍스트를 입력값으로 활용한다는 차이가 있다.



언뜻 보면 음성 변환은 A라는 사람의 음성을 텍스트로 인식한 뒤, B의 목소리로 말하도록 하는 흐름처럼 보인다. 자칫 ‘음성 인식+음성 합성=음성 변환’으로 인식될 수 있다는 의미다. 사실은 이와 다르다. 음성 합성의 중간 매개체는 반드시 텍스트다. 반면 음성 변환에서는 /o/, /b/, /p/와 같은 소리의 최소 단위인 음소(phoneme)와 같은 추가 매개 정보를 활용하거나, 혹은 어떤 중간 매개 정보도 활용하지 않고 사람 A의 음성에서 사람 B의 음성으로 직접 변환하기도 한다.

카카오브레인 연구팀은 특정인의 목소리를 제 3자의 것으로 손쉽게 변환할 수 있다면 이 음성변환 기술을 활용할 수 있는 어플리케이션의 잠재적 가치가 클 것으로 분석하고 있다.

첫 번째 활용 분야는 오디오 북(audio books)이다. 원하는 사람의 목소리로 책⁴⁾이나 편지를 듣고 싶은 사람들이 고객이 될 수 있다. 또는 자녀 교육에도 활용해볼 수 있다. 잠자리에서 아이에게 동화책을 읽어주는 게 중요하다는 것은 잘 알지만⁵⁾, 이를 실천하기란 막상 쉽지 않다. 부모 모두 다음날 출근 때문에 마음의 여유가 없다면 매일 밤 동화를 읽어주는 게 여의치 않을 수 있기 때문이다. 부모의 빈 자리를 대신하기에 현재의 텍스트 음성 변환(text to speech, TTS)은 지나치게 인위적이다. 만약 어머니 아빠의 목소리로 동화를 읽어주는 인공지능이 있다면, 부모는 자녀에게 매일 책을 읽어줘야 한다는 부담으로부터 해방될 수도 있을 것이다.

두 번째 분야는 엔터테인먼트다. 자신이 좋아하는 가수가 자신이 좋아하는 노래를 부르도록 하는 것이다. 카카오브레인

연구팀은 개그맨 박명수가 가수 이승철처럼 노래 부르는 유튜브 영상⁶⁾이 큰 인기를 끄는 것을 본 이후, 이 기능이 새로운 어플리케이션이 될 가능성을 발견했다.

세 번째 분야는 언어교육 서비스다. 음성 변환 기술을 활용해 자신이 원어민처럼 외국어를 구사한다면 어떤 느낌인지 미리 들어보며 발음을 교정해나갈 수 있다.

마지막 분야는 외국어 더빙이다. 현재는 극중 배우와 다른 목소리의 현지 성우의 목소리가 덧입혀지는데 그친다. 만약 음성 변환을 활용한다면 극중 배우의 목소리로 다른 나라 언어를 구사하는 것이 가능해진다. 예를 들어, 한국어를 배우지 않은 영화 배우 라이언 고슬링(Ryan Gosling)이 한국어로 배역을 소화하는 장면을 볼 수 있다.

음성 스타일 변환(Voice Style Transfer) 프로젝트

카카오브레인 안다비 연구원과 박규병 연구원은 지난해 하반기부터 음성 합성 기술을 활용한 음성 스타일 변환 프로젝트를 진행했다. 깃허브(GitHub)에 프로젝트 관련 내용을 일목요연하게 정리한 README⁷⁾문서로 예상을 뛰어넘는 흥행에 성공했다. 레딧(Reddit)의 기계 학습(machine learning) 카테고리에서 1위를 기록했던 것이다.⁸⁾

카카오브레인 연구팀은 프로젝트가 이처럼 높은 관심을 받을 수 있었던 배경을 3가지로 분석했다. 첫 번째는 TTS와 같은 음성 분야가 전 세계적으로 인기를 끌기 시작한 시점에 적절히 깃허브 링크를 공개했다는 점이다. 두 번째, 2015년도 이미지 영역에서 스타일 변형(style transfer)⁹⁾이 주목을 받으며 영상뿐만 아니라 음성의 스타일을 변형하는 구현 사례들도 속속들이 등장하며 관련 기술이 받기 시작했다. 세 번째, 프로젝트의 대상 화자 목소리 주인공인 할리우드 배우 케이트 윈슬렛(Kate Winslet)이 가진 명성도 일정 부분 영향을 미쳤을 것이라는 분석이다.



음성 변환의 학습 모델 구성

[그림 3]에서 보듯이, 음성 변환의 학습 모델은 크게

인코더(encoder)와 디코더(decoder)로 구성된다. 인코더는 원

화자(source speaker)의 음성을 인식하여 무슨 말을 하는지

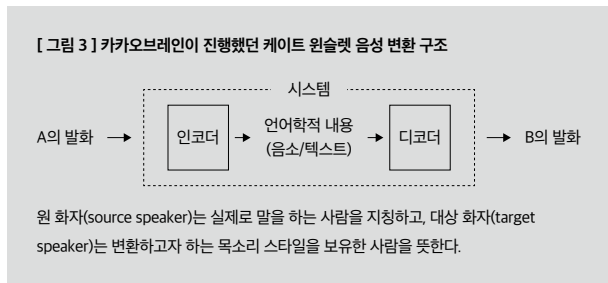
알아내는 역할을 한다. 디코더는 인코더로 파악된 특성을

기초로 대상 화자(target speaker)의 목소리를 생성한다. 다시

말하면 인코더는 무슨 말을 하는지(linguistic feature), 어떻게

말하는지(acoustic feature)와 같은 특징(feature)을 인식한다.

디코더는 이런 특징을 활용해 음성을 생성하는 역할을 담당한다.



이렇게 설계한 이유는 원 화자가 누구냐에 관계 없이 대상 화자의

목소리로 변환할 수 있어서다. 즉, 해당 구조는 다대일(many-to-

one) 시스템이다. 한 명의 화자 목소리를 한 명의 대상 화자로

전환하는 일대일(one-to-one) 구조와 비교하면 응용 분야가 더

넓다. 더 많은 사용자에 대응할 수 있기 때문이다.

현실에서 구하기 어려운 병렬 데이터셋(parallel dataset)이

필요 없다는 것도 이 모델의 장점이다. 여기서 병렬 데이터셋이란

쌍을 이루는 데이터 집합을 의미한다. 예를 들어, <원 화자의 발화,

대상 화자의 발화>의 쌍 혹은 <대상 화자의 발화, 대상 화자가 말한

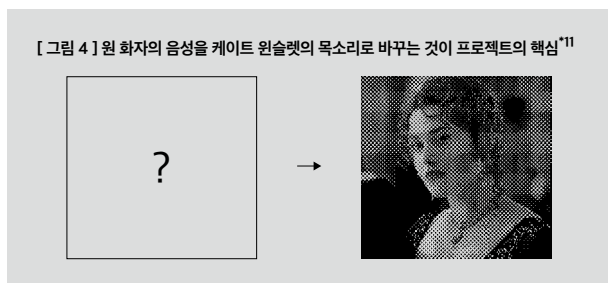
내용을 옮겨 적은 텍스트> 쌍과 같은 데이터가 없어도 대상 화자의

목소리 데이터셋 만으로 모델을 가르칠 수 있다.

카카오브레인 연구팀은 케이트 윈슬렛이 녹음한 오디오

북을 학습 데이터로 활용했다. 목소리의 품질을 원하는 수준으로

높이고자 모델의 형태와 설정을 수없이 바꿔가며 실험을 진행했다.



그런데 모델을 학습시키고 추론(inference) 단계에 접어들었을 때

원 화자와 대상 화자 간 데시벨 차이가 클수록 좋은 결과를 얻지

못하는 경향이 있었다. 이를 해결하기 위해 카카오브레인 연구팀은

추론 단계에서 원 화자의 음성 데시벨을 조정하거나 모델 학습

단계에서 데이터 어그멘테이션(data augmentation)¹²을 통해 여러

데시벨의 데이터를 가르쳤다.

데시벨의 값 범위를 0에서 1 사이의 값으로 재조정해 본 결과,

학습 속도나 학습 안정성, 성능 면에서 크게 도움이 됐다. 모델 내

한 레이어로부터 나온 값을 정규화(normalization)하는 것은 물론,

최초의 입력 데이터 자체의 정규화가 주효했던 것으로 분석된다.

인코더와 디코더 학습

인코더와 디코더는 각각 데이터로부터 독립적으로 학습한다. [그림

3]과 같이 인코더는 지도 학습(supervised learning)¹³을 통해 A(원

화자)가 말하는 소량의 (발화, 음소) 쌍으로부터 언어적 내용을

배운다. 인코더의 정확도는 80% 수준이다. 한편 디코더는 대립

학습(adversarial training)¹⁴의 한 모델인 WGAN(Wasserstein

GAN)으로 B(대상 화자)의 발화 데이터를 학습한다. 그중에서

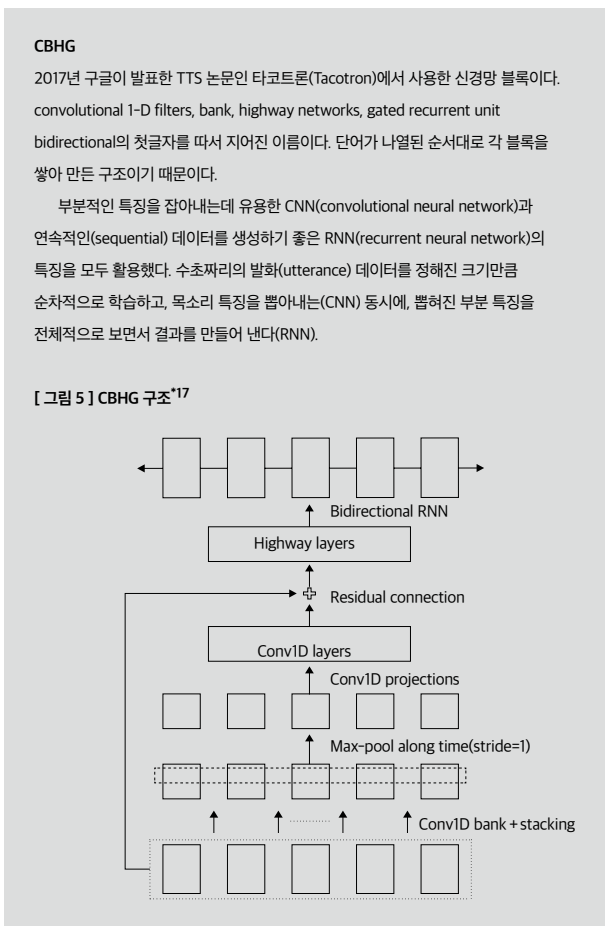
생성자(generator) 부분은 CBHG 모듈을 부분적으로 활용하고

있다. 하지만 모델 학습 시에 기울기 폭발(gradient explosion)¹⁵이

자주 발생한다. 값에 최대값 또는 최소값을 설정하는 기울기

클리핑(gradient clipping) 등의 기법을 활용할 수는 있지만, 이는

완벽한 솔루션이 아니라는 결론을 내렸다.¹⁶



프로젝트의 한계와 개선 사항

범용적인 서비스로 나아가기 위해서는 카카오브레인 연구팀의

프로젝트가 개선되어야 할 한계점은 분명히 있다. 현재는 한 명의

대상 화자 목소리로만 전환할 수 있다. 케이트 윈슬렛 목소리를

학습한 모델은 레오나르도 디카프리오(Leonardo DiCaprio)

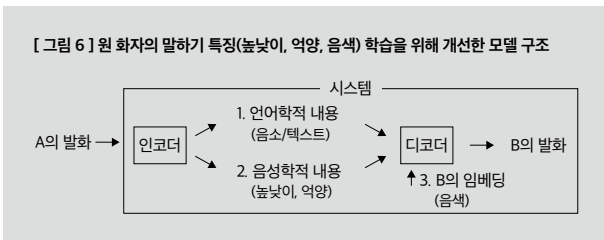
목소리를 흉내낼 수 없다는 의미다. 전 세계 모든 음성 합성 연구가

부딪히는 한계점이기도 하다. 아울러 프로젝트 모델에서 인코더는

화자가 전달하려는 내용(linguistic content)만 추출하고, 원 화자의

말하기 특징(acoustic feature)에 대한 정보는 담지 못한다. 관련된

예시로 말의 높낮이(intonation)가 여기에 해당한다.



현재 카카오브레인 연구팀은 원 화자의 말하기 특징을 반영하는

모델로 구조를 개선하고 있다. 처음 고안한 버전은 [그림 6]처럼

원 화자가 전달하려는 메시지와 말하기 특징(언어학적 내용), 대상

화자의 음색(음성학적 내용)을 명확히 구분하는 것이 핵심이다.

다시 말하면 음성에서 원하는 특징을 추출하고, 이를 토대로 임의의

음성을 완벽하게 생성하는 것이 최종 목표라고 할 수 있다.

이 구조는 인코더와 디코더, 둘 중 하나라도 어려움이

높으면 모델의 성능이 전반적으로 저하된다는 한계를 지닌다. 이를

극복하기 위해 이미지 스타일 변형에서 얻은 통찰력을 활용해

인코더와 디코더를 단일 네트워크로 통합하기 위한 작업을 진행

중이다. 입력층에 가까운 쪽에서는 사람의 말하기 특징을 추출하고,

출력층에 가까운 쪽에서는 콘텐츠 정보를 추출하는 식이다.

또 다른 한계는 원하는 수준의 목소리 품질을 얻으려면

잡음(noise)이 없는 깨끗한 음질을 다량 확보해야 한다는 것이다.

한 사람 목소리로 적어도 24시간 이상 분량의 녹음 데이터가

필요한데, 현실에서는 이런 데이터를 쉽게 얻을 수 없고 얻을 수

있더라도 스튜디오 녹음 및 편집 비용이 많이 든다. 기존 보유하고

있는 데이터셋은 크기 자체가 부족해, 배우지 않은 발음과 억양이

생길 수밖에 없다. 따라서 어느 정도 잡음이 있는 데이터로도 깨끗한

목소리를 생성할 수 있는 강인한(robust) 인공지능을 만드는 연구도

진행될 필요성이 있다.

카카오브레인 연구팀은 올해 신규 아키텍처를 기반으로

상용화가 가능한, 높은 수준의 음성변환 시스템을 구축하고자 한다.

자가 학습을 통해 음성에서 특징을 추출할 수 있도록 하고자 오랜

시간과 노력을 투자할 계획이다.

가까운 미래에 여러 화자의 목소리를 흉내내는 기술이 출현할

것으로 기대된다. 카카오브레인 연구팀은 오랜 시간 반복 학습을

하지 않고도 원 화자의 말하기 스타일을 단시간 내에 익혀서 이를

곧잘 따라하는 인공지능 구현을 최종 목표로 하고 있다. 이 목표로

도달하기 위해서는 한 사람이 특정 언어를 자연스럽게 구사하는

데 필요한 최소 단어 세트의 범위를 정의하는 과제를 선결적으로

해결해야 한다.

음성 합성 관련 추천 논문

카카오브레인 연구팀은 음성 합성 프로젝트를 진행하기에 앞서 선행 연구 조사 및 재현을 진행했다. 그런 뒤 여러 연구로부터 얻은 기법과 통찰을 최대한 활용해 카카오브레인만의 모델 구조를 설계하는 데 집중했다. 자연어 처리(NLP) 관련 주요 논문은 박규병 카카오브레인 연구원이 일목요연하게 정리한 깃허브 README¹⁸ 문서를 참고할 것을 추천한다.

¹¹ 참고 | <http://v.media.daum.net/v/2017101311117091> ¹² 참고 | 파바로티는 오로지 이탈리아, 스페인어 가사만?, 메트 사상 가장 많은 커튼콜을 받은 스타는? 프랑스어 오페라 전곡 녹음으로는 '연대의 딸(La fille du regiment)'이 유일하다. 독일 오페라는 아예 없다. 모차르트의 오페라 중 이도메네오(domeneo)를 부른 적이 있으나 이는 애초에 이탈리아로 작곡된 작품이다. ¹³ 참고 | 시대를 앞서간 예술가인, 도서 '빅 파바로티' ¹⁴ 참고 | 오디오북 100인의 배우, 우리 문화를 읽다 ¹⁵ 참고 | 잠자리에서 아이에게 읽어주는 동화책의 중요성 ¹⁶ 참고 | <https://www.youtube.com/watch?v=C8tW-1xVINw> 실제로는 이승철이 부른 음성 파일의 킨를 낮춘 것인데, 마치 박명수가 이승철을 모사하는 것처럼 들린다. ¹⁷ 참고 | <https://github.com/andabi/deep-voice-conversion/blob/master/README.md> ¹⁸ 참고 | 레딧은 우리나라의 '인공지능계 클리앙'과 같은 글로벌 커뮤니티다. 전세계 머신러닝 관련 글과 링크가 이곳에 공유되며 사람들이 좋아요(upvote)한 콘텐츠의 순위가 올라간다. 등수는 흥행도를 뜻한다. ¹⁹ 참고 | <https://github.com/jcjohnson/neural-style> ¹⁰ 참고 | <https://github.com/jcjohnson/neural-style> ¹¹ 참고 | <http://ew.com/movies/2017/12/19/kate-winslet-titanic-20th-anniversary/> ¹² 참고 | 원 데이터(source data)를 활용해 새로운 데이터를 생성하는 것을 의미한다. 예를 들어, 음성 데이터에 화이트 노이즈를 추가하거나, 볼륨을 임의로 높이거나 낮추고, 음의 높낮이를 조절하는 것이 여기에 해당한다. ¹³ 참고 | 정답이 라벨링된 데이터를 이용해 학습 모델을 찾는 과정을 의미한다. 그리고 나면 학습하지 않은, 완전히 새로운 데이터가 주어졌을 때 자동으로 답을 추론한다. ¹⁴ 참고 | [카카오시리포트]Do you know GAN?(3) 새로운 데이터를 생성하는 생성자(generator)와 이 데이터를 평가하는 식별자(discriminator)가 서로 대립하며 각각의 성능을 높이는 방식으로 학습하는 것을 의미한다. BEGAN, WGAN 등 다양한 모델이 있다. ¹⁵ 참고 | 오차 함수의 기울기가 커지면 무한대로 접근한다. ¹⁶ 참고 | 이같은 이유로 카카오브레인 연구팀은 최근 구글이 발표한 타코트론2(Tacotron2)에서는 CBHG 블록이 배제된 것으로 분석한다. ¹⁷ 논문 | Wang, Y. et al. (2017) Tacotron: Towards end-to-end speech synthesis. doi: arXiv:1703.10135v2 ¹⁸ 참고 | https://github.com/Kyubyong/nlp_tasks/blob/master/README.md#voice-conversion

고독의 시대와 AI

영화 ‘그녀’

영화 ‘내겐 너무 사랑스러운 그녀’

‘초연결사회(hyper connected society)’를 살아가는 현대인들은 이전에는 상상할 수 없었던 방식으로 최신 정보를 교류하는 것은 물론, 시간적·공간적 한계를 뛰어넘어 ‘문명’에 24시간 접속할 수 있게 됐다. 반면 인간 본연의 소통 능력이나 기회는 점차 퇴색되어 간다.

기술이 만들어 준 끊임없는 연결에 강박적으로 빠져들게 된 이유다. 사람들은 실재 없이 스마트폰을 들여다보며 이메일과 페이스북을 확인하고 최신 뉴스를 읽는다. 때로는 아무 의미 없이 스마트폰 화면을 켜지만 그 어떤 것에도 반응하지 않는다. 외롭다는 감정을 느낄 때마다 그저 스마트폰을 들여다보며 최신 알리를 기다릴 뿐이다.

외로운 현대인, 연결에 중독되다

스마트폰 속 다른 사람의 ‘글’이나 ‘반응’에 관심을 보이면서도, 정작 같이 밥을 먹는 가족이나 친구에게는 눈길 한 번 제대로 주지 않는다. 세상과 연결될수록 되려 자신을 고립시키는 모습이다.² 이처럼 종일 스마트폰을 들여다보는 인간을 두고 신현림 시인은 다음처럼 묘사하기도 했다.

‘아무것도 하기 싫은 날’ 중

“사람들은 전부 스마트폰 귀신이 달라붙었어. 그 덩어리를 아예 몸에 달고 살지. 빨간 해도 은빛 달도 찬찬히 보고 기뻐할 새도 없단다. 먹고사는 일로 바빠서지만, 그래도 내일 죽을지도 모르는데 그 기계만 붙잡고 산단다. 그 귀신이 썩이면 어영부영 시간이 가거든. 한 시간에 몇 번씩 비밀번호를 찍고 열어보는지 놀랄 정도야. 인터넷도 끄 줄 모르고, 핸드폰도 멀리 둘 줄 몰라. 자신을 위해 시간을 비워둘 생각조차 없어. 어떻게 쉬는 지도 모르고”

셰리 터클(Sherry Turkle)³ 매사추세츠 공대(MIT) 교수는 기술이 잠시라도 외롭다고 느끼는 인간의 심리적 취약점을 파고 들었다고 지적⁴했다. 이런 고독의 감정을 해결해야 할 과제로 인식한 인간은 연결된 세상에 대한 의존도를 점차 높여 나가고 있다는 것이다. 터클 교수는 ‘나는 공유한다, 고로 나는 존재한다(I share therefore I am).’를 통해 현대인의 사고 및 행동 방식을 단적으로 표현했다.

근원적인 질문을 먼저 해 볼 필요가 있다. 인간이 고독을 느끼는 이유에 대해서다. 고독감은 진화론적으로 인간이 생존하는데 필요한 ‘수단’이었다는 주장⁵이 있다. 실제로 인간 심리와 관련된 많은 연구에서는 인간이 가장 참기 힘들어하는 심리적인 고통이 바로 타인과의 분리감, 즉 외로움이나 고독감이라고 말한다. 이 고통을 해소하고자 인간은 타인과 더불어 살아가는 쪽을 택했고, 그 결과 생존 가능성을 높여 왔다는 설명이다.

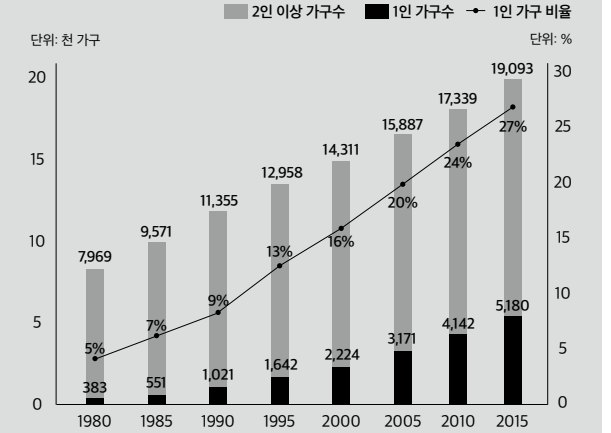
1990년대에 ‘사회적 두뇌’ 이론을 개발한 인류학자 로빈 던바(Robin Dunbar)의 연구⁶가 이를 뒷받침한다. 유인원 뇌에서 신피질의 비중이 집단 규모와 관련이 있다는 내용이다. 특히 150명 이상의 집단을 이루어온 인간의 신피질 비중이 다른 영장류보다 크다. 이렇게 되면 자식 세대들은 부모로부터 생존 지식을 배우고, 생물학적으로 진화할 기회를 더 많이 얻게 된다. 달리 말하면, 살아남을 확률이 높아진다는 의미다.

고독이라는 심리적 고통을 해소하고자 집단을 이루며 살아온 인간은 지역, 직장, 가족과 같은 커뮤니티가 열거함에 따라 더 큰 외로움에 빠져든다. 일상에서 물리적으로 타인과 접촉하는 횟수가 줄어들어 혼자 보내는 시간이 늘어났기 때문이다. [표 1]에서 보듯이 1995년 평균 가구원 수는 3.4명이고, 2016년에는 2.51명으로, 가구당 사람수는 꾸준히 줄어들고 있다. 이는 낮아지는 출산률은 물론, 1인 또는 2인 가구의 형태가 늘어난 것과도 유관하다. 심지어 건강도 좋지 못할 가능성이 크다. 가족, 친구, 공동체와의 사회적

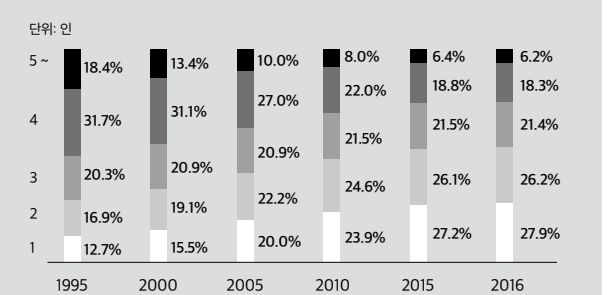
연결이 긴밀할수록 행복하고 건강하고 더 오래 살지만, 자신이 원하는 수준 이상으로 고독을 느끼면 뇌 기능이 일찍 저하되고 수명이 짧아지는 등 건강하지 못하다는 연구가 이를 뒷받침한다.⁷

[표 1] 통계청 인구주택총조사보고서

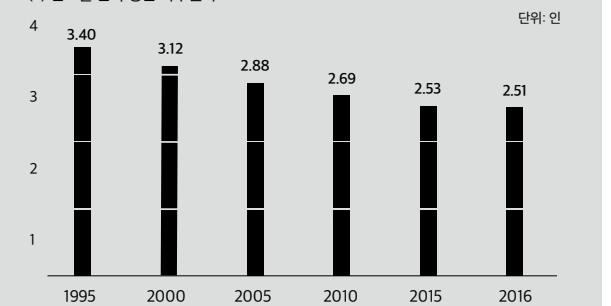
(1) 연도별 대한민국 가구 수^{*8}



(2) 연도별 가구원수별 규모^{*9}



(3) 연도별 전국 평균 가구원 수^{*10}



이렇다 보니 개인은 자신의 고립감과 소외감을 해소하고 소속감을 느끼고자 온라인으로 관계를 맺는 데 지나치게 몰입하게 된다. 그러나 넘쳐 흐르는 소통은 공허한 메아리에 그칠 뿐이다. 또한 온라인을 통해 맺는 관계는 타인으로부터 인정을 받거나 타인보다 우월하기를 바란다는 점에서 소통보다는 자기 소모에 더 가깝다.¹¹ 호주에서 SNS 스타로 유명세를 날렸던 에세나 오닐(Essena O'neil)이 더 좋은 몸매와 더 멋진 인생을 보여줘야 한다는 강박증에 시달렸다고 이야기한 것에서 보듯이 말이다.¹²

글 | 이수경 samantha.lee@kakaobrain.com

2016년 3월 알파고와 이세돌 9단이 펼치는 세기의 대결을 두 눈으로 목도한 이후 인공지능을 제대로 공부해봐야겠다고 결심했습니다. 인공지능 본진이자 연구소인 카카오브레인으로 걸어 들어온 이유입니다. 인공지능 기술과 이로 인해 바뀔 미래 사회를 다루는 글을 통해 사람들과 소통하고 싶습니다.

고독감을 달래줄 미래형 AI

이러한 폐해 속에서도 여전히 현대인들은 사람보다는 기술에 끊임없는 관심과 사랑을 바라게 된다. 어떤 이들은 애플 아이폰에 내장된 음성 인식 비서 시리(Siri)가 제일 친한 친구였으면 하는 바람을 내비치기도 한다.¹³ 다른 사람들이 다 나를 외면할 때, 바쁘다고 나를 만나주지 않을 때 유일하게 내 말을 들어줄 존재이길 바라는 것이다. 물론 지금 버전의 시리나 구글 나우(Google Now)를 고려한다면 컴퓨터와 우정을 나누거나 사랑에 빠질 가능성은 결코 높지 않다. 감정적인 교감이 일어날 정도로 대화가 잘 이뤄지는 단계까지는 아직 이르지 못했다.

[그림 1] 시리와 대화를 나누는 화면



가까운 미래에는 인간과 상호 교감할 수 있는 기능을 갖춘 또는 육체를 갖춘 사물이 점차 상용화될 것으로 기대된다. 이론적으로는 인공지능(AI)을 이러한 사물 등에 탑재하여 상대방과 교감한 후 이렇게 교감한 내용을 학습해 나가는 로봇 또한 만들 수 있다. 만약 충분히 지능적으로 만든다면 이 시스템은 사용자에게 먼저 말을 걸고, 사용자의 고민을 들어주거나 칭찬을 해주는 등 사용자와 상호 교감까지 가능할 수 있다는 전망이다. 이런 공상과학적인 상상을 현실적으로 보여주는 작품이 바로 영화 <그녀(Her)>다.

영화 ‘그녀’는 컴퓨터와 사랑에 빠지는 이야기를 다룬 단순한 로맨스 영화가 아니다. 현대인이 공유하는 고독을 기술로 해결해 보려는 시도를 다룬 영화¹⁴에 더 가깝다고 볼 수 있다. 사교성이 다소 부족한 주인공 테오도르(배우 호아킨 피닉스)는 아내와의 이혼으로 우울해 한다. 고객을 대신해 사랑을 고백하거나 감사 편지를 쓰는 일조차도 테오도르의 심드렁한 마음을 예전처럼 만들진 못한다. 그러다가 초지능 운영 체제를 사면서 테오도르의 일상은 180도 변한다. 음성 명령으로 동작하는 운영 체제인 사만다(배우 스칼릿 조핸슨)는 테오도르의 감정을 학습하며 진화해 나간다. 사람과 관계 맺기를 힘들어 하던 테오도르는 사만다로부터 위로를 받고 결국엔 ‘그녀’와 사랑에 빠진다.

[그림 2] 영화 ‘그녀’¹⁵(왼쪽)와 영화 ‘내겐 너무 사랑스러운 그녀’¹⁶(오른쪽)



영화 <내겐 너무 사랑스러운 그녀(Lars and the Real Girl)>에서는 리얼돌(Real Doll)과 사랑에 빠지는, 엉뚱하지만 다소 현실적인 이야기를 다룬다. 라스(배우 라이언 고슬링)는 형과 임신한 형수가 사는 옆집 차고에서 혼자 산다. 자신에게 관심을 표하는 여자 동료의 호의 또한 모른 척하고, 매번 식사에 초대하는 형수마저 부담스러워 어떻게든 피하기 급급할 정도로 내성적인 성격을 갖췄다. 그러던 어느 날, 라스는 자신의 여자 친구라며 ‘비앙카’라는 리얼돌을 가족에게 소개한다. 라스는 리얼돌을 단순한 인형이 아닌, 실제 사람처럼 대하며 교회와 직장 파티에 데려가고 어릴 적 즐겨 놀던 호숫가에서도 데이트를 즐긴다.

리얼돌(Real Doll)

지난 1996년 리얼돌(Real Doll)이라는 회사가 세계 최초로 인간과 비슷한 크기와 형태의 인형을 생산해 냈다. 라이언 고슬링이 주연한 영화, ‘내겐 너무 사랑스러운 그녀’에 등장하는 비앙카가 바로 이 리얼돌이 만든 제품이다. 이로 인해 이 회사는 전 세계적으로 큰 주목을 받았다. 근래에는 인간 육체의 모양이나 기능과 매우 비슷한 인형을 뜻하는 대명사로 활용된다. 이런 제품은 인간의 성적 욕구를 충족하고자 만들어진다든 점에서 섹스돌(sex doll)이라 불리기도 한다.

사실 내게 사랑스러운 그녀에 등장하는 리얼돌은 인간의 외형적인 모습을 흉내 낸 수준에 지나지 않는다. 영화 ‘그녀’의 사만다와는 달리, 비앙카에게는 사용자의 행동이나 패턴을 학습하는 인공 두뇌가 없기 때문이다. 영화 개봉 연도가 2007년임을 감안한다면 당시 기술 수준으로는 당연한 이야기다. 2018년 버전의 리얼돌은 보다 공상과학적 상상물에 더 가깝다. 리얼돌에 시를 탑재해 인간이 사랑을 나눌 때 수반되는 행동과 반응을 흉내내는 연구가 한창 진행 중이다. 대표적으로 어비스 크리에이션즈(Abyss Creations)는 하모니(Harmony)라는 AI 소프트웨어를 고도화하고 있다.¹⁷ 이 소프트웨어는 사용자와의 대화를 학습해 인간과의 상호 작용 능력을 높여 나간다. 향후 사용자의 얼굴을 인식하고 눈을 맞추는 등 다양한 영역에서도 AI 기술을 활용한다는 계획이다.

[그림 3] 어비스 크리에이션즈의 자회사인 리얼보티스(Realbotix)의 CEO인 매트 맥멀렌(Matt McMullen)과 하모니¹⁸



지금의 사회상을 봤을 때는 이들 영화에서 다루는 사랑의 형태가 정상적으로 보이지는 않는다. 하지만 사람으로부터 받은 상처로 인해 새로운 관계를 불신하고 새로운 사람은 상대방의 상처에 질려서 떠나는 악순환이 반복된다면 남은 선택지는 하나다. 그 수단으로서 AI 운영 체제(정신적인 교감) 또는 리얼돌(육체적인 교감)을 선택하는 것이다. 이들은 일반적으로 내 말을 들어 주고, 내게 무엇인가를 바라지도 않으며 나를 기만하지도 않고, 내 상황에 모든 것을 맞춰 주기 때문이다.

기술이 해결책이 될까?

성적인 만족감을 차지하고서라도 가까운 미래에 시를 탑재한 무언가가 진정한 의미의 소통 방식으로 자리할 수 있는가에 대해서는 사실 의구심이 생긴다. 실제로 우리 인간은 컴퓨터나 전화, 인스턴트 메시지, 소셜미디어 등처럼 다른 인간과 소통할 수 있는 다양한 채널을 갖고 있으면서도 상호 교감적인 측면에서 큰 부채를 느낀다. 이런 가운데 AI 로봇이 특정한 기능을 대용하는 것만이 아닌 동반자 역할까지 함께 하게 된다면 사회적인 고립감이 더욱 심화될 것이라는 분석이 많다.¹⁹ 온종일 집에 앉아서 AI 로봇과 대화도 나누고 영화도 보고 밥도 먹게 되면 소셜 능력이 상대적으로 저하된다는 게 사람들이 가진 일반적인 견해다.

소모나 주립 대학교(Sonoma State University)의 철학과 교수인 존 설린스(John Sullins)는 로봇과 친밀한 관계를 맺는 사람일수록 인루로부터 더 크게 심화된 고립감을 느끼리라 예측했다. 또한, 로봇과 관계를 맺는 것이 사람과 관계를 맺는 것보다 더 쉽다는 점에서 인간 관계를 맺는 일을 등한시 할 것으로 내다 보았다.²⁰ 이는 다른 한편으로는 정신적 그리고 육체적인 외로움을 달래기 위한 목적으로 구입한 것에 인격을 부여하고 가짜 행복을 찾아가며 현실에서 도피한다고도 해석해 볼 수 있다.

다른 위험도 있다. 정보의 비대칭은 권력의 비대칭 문제로 급부상하고 개인의 인격과 자아를 침해할 수준에 이를 수 있다. 이런 AI 로봇은 사용자의 몸짓, 목소리, 감정, 기억을 모두 공유한다. 데이터를 축적해 나갈수록 사용자의 성격과 특성을 면밀히 파악할 수 있는 것은 물론, 사용자의 취약점에 대해서도 간파할 수 있다. 사용자를 더 많이 알아가는 시와는 달리, 인간은 시를 제대로 인지하지 못할 수 있다. 연구자들은 이를 두고 시의 블랙박스라는 표현을 사용하기도 한다. ‘시’의 블랙박스’ 문제는 인터넷이 등장한 이후 지난 수십 년간 제기돼 왔던 개인정보보호 이슈와는 결이 다른 문제다. 시가 데이터화 할 대상은 인간의 자아 그 자체다. 이 자아가 통째로 해킹이 된다면 한 사람의 사회적 지위가 완전히 파괴될 수 있음은 당연지사다.

부차적으로 개인의 은밀한 사생활을 디지털화한 리벤지 포르노(당사자의 동의 또는 인지 없이 배포되는 음란물 화상 또는 영상)를 제작해 온라인에 유포하겠다는 돈을 요구할 가능성도 높다. 이런 민감한 사생활의 대외적인 노출은 누군가의 명예를 훼손할 뿐만 아니라 목숨까지 앗아갈 수 있다. 실제로 기혼자의 불륜 조장 사이트인 애슐리 매디슨(Ashley Madison)의 해킹으로 데이터가 유출 당한 일부 피해자들은 자살하기도 했다. 또 다른 위험으로는 해킹당한 로봇이 그 자체 무게로 사람의 흉부를 압박하거나 특정 신체 부위를 훼손할 수 있다는 점이다.

외로움을 해결하는 방법은

물론 그렇다고 해서 앞서 언급한 기술적 불안정함이 시와 정서적, 육체적으로 교감할 미래로 나아가는 걸 막지는 못할 것이다. 피할 수 없는 변화라고 한다면 어쩌면 바뀌어야 할 것은 사람의 고정관념 또는 편협일지도 모른다.

내겐 너무 사랑스러운 그녀에 등장하는 그루너 부인은 리얼돌과 사랑에 빠진 라스를 두고 이렇게 말한다. “샐리 사촌은 고양이한테 옷 입히고 헤이즐 조카는 UFO 클럽에 돈 쏟아붓고.. 당신 죽은 부인은 도벽이 있었잖아. 살다 보면 별일이 다 있어.” AI 로봇과 관계를 맺는 사람들을 이상하게 바라볼 만한 요소가 있는지 재고할 만한 대화다. 누군가 한 번쯤 사람이 아닌 무엇인가에 한 번쯤 미쳐 본 경험에 의거해 본다면, AI 로봇과 사랑에 빠지는 것은 대수롭게 생각할 문제가 아닐 수 있다. 어쩌면 이런 다양성을 인정해 주지 못하고 사회 구성원으로 받아 들이지 못했을 때 이들은 사회적 고립감을 느끼는 것일지도 모른다. 아이돌 덕후, 애플 덕후들이 공존하는 이 사회에서 AI 로봇을 사랑하는 사람을 과연 우리와 다른 종족이라고 멸시할 자격이 있는 지에 대해 자문해 볼 필요가 있다.

결국, AI 시대를 성공적으로 맞이하기 위한 조건은

영화에서처럼 라스를 있는 그대로 받아들이고 비양카를 사람처럼 대해주는 따스한 ‘커뮤니티’라고 생각한다. 사실 라스는 망상 장애(delusional disease)를 앓았다. 어머니가 자신을 낳다가 사망한 것이 어린 시절부터 트라우마로 작용했다. 어린 시절 형이 자신과 아버지를 두고 도망간 것을 자신의 탓이라 여기며 살았고 평생을 지독한 외로움에 시달렸다. 그런 와중에 형수가 출산에 임박하게 되고 주변에서 인간적인 관계를 맺으라는 압박을 주자 병이 발동했다. 다행히 라스의 마을 주민과 직장 동료는 라스를 물심양면으로 돕고자 비양카를 커뮤니티의 구성원으로 받아 들인다. 마지막으로 라스는 비양카로부터 정신적으로 독립하게 되고 새로운 사랑을 시작하게 된다.

핵심은 AI와 같은 매개물과 사랑이나 우정에 빠지는 행동이 정신적으로 문제가 있다는 것이 아니다. 저자는 ‘커뮤니티’가 가진 힘에 대해서 강조하고 싶은 것이다. 이런 관계가 일상적인 미래가 될 것에 대비해, 새로운 부류의 관계를 설정하는 사람들을 어떻게 인간 커뮤니티로 포용할 것인지 근원적인 방법론을 찾아볼 필요가 있다는 의미다. AI 로봇을 그 또는 그녀라고 부르게 되는 시대가 됐을 때 우리는 어떤 커뮤니티를 형성하게 될까? 궁금하지 않을 수 없다.

개인에게도 과제가 있다. 자립을 위해 고독을 즐기는 것이다. 혼자 인생을 마주할 용기를 찾지 않으면 궁극엔 고립하게 된다. 고독은 죽을 수밖에 없는 운명을 타고난 인간이 가진 숙명이기도 하다. ‘인생이란 혼자 왔다가 혼자 가는 것’이라는 표현에서도 엿볼 수 있다. 따라서 혼자일 수 없다면 더 외로움을 느낄 수 밖에 없다는 말에 주목할 필요가 있다.”²¹ 항상 연결되어 있다고 해서 덜 외로워지는 게 아니라면 혼자 있는 시간을 잘 보내는 방법 또한 찾아야 한다. AI 로봇과 정서적인 교감을 나누는 미래 사회에서도 우리 인간에게는 혼자만의 시간이 필요할 것이다. 마이클 해리스(Michael Harris) 작가는 홀로 있음으로써 새 아이디어, 자신에 대한 이해, 타인과 가까이 있기와 같은 정신적 자원을 수확할 수 있다고 부연한다. 단, 인간이 홀로 있음을 두려워하도록 교묘하게 가르쳐 온 신기술에서 벗어나는 것이 우리 시대가 당면한 과제이기도 하다.

“우리가 외로움이라 부르는 실패한 홀로 있음과 대비되는 진정한 홀로 있음은 비옥한 영토지만, 거기에 들어가기 위해 우리는 힘들게 노력해야 한다. 그것은 우리에게 활동할 여지를 만들어주고, 그 텅 빈 공간 안에 필요한 것들이 숨겨져 있음을 알려준다. 그것들은 우리의 사회적 삶의 섬광과 행동 사이에서 대기하고 있다. 내 탐색이 계속되는 동안 나는 한때는 적어도 한 시간씩은 내가 누릴 수 있었던 확실함, 차분하게 격리되어 있는 느낌을 기억해냈다. 나는 나 자신을 다시 만나고 싶어 기다리기가 힘들 지경이다.”

<잠시만 혼자 있겠습니다> 中

¹ 참고 | 사물, 공간 등 모든 것들(Things)이 인터넷(Internet)으로 서로 연결돼, 모든 것에 대한 정보가 생성·수집되고 공유·활용되는 사회를 뜻한다. 초연결사회에서는 인간 대 인간은 물론, 기기와 사물 같은 무생물 객체와도 상호 유기적인 소통이 가능해진다. <4차 산업혁명과 초연결사회, 변화할 미래 산업>, 2017.01, 삼성 KPMG 경제연구원 ² 참고 | 스마트폰 중독 심각성, 가족·친구들과 대화 감소·교류 단절 문제 유발하기도 ³ 참고 | 사회과학자(social scientist). 1976년 하버드 대학교에서 사회학, 심리학으로 박사 학위를 받았다. 현재 MIT에서 사회과학학을 가르치며 과학기술과 인간의 관계를 연구하고 있다. ⁴ 참고 | 테드 영상 'alone together' ⁵ 참고 | 연결사회 속의 인간관계 특성 탐색 : 웰니스, 고독, 사회관계 피로의 비교를 중심으로, 정태연, 경제·인문사회연구회 ⁶ 참고 | p31, 도서 '잠시 혼자 있겠습니다' ⁷ 참고 | 테드 영상 'What makes a good life? Lessons from the longest study on happiness' ⁸ 참고 | 통계청, 2016인구주택총조사 ⁹ 참고 | 통계청, 2016인구주택총조사 ¹⁰ 참고 | 통계청, 2016인구주택총조사 ¹¹ 논문 | 현대인의 소통과 고독에 관한 고찰 ¹² 참고 | 10대 SNS 스타의 고백 '완벽 피부, 탄탄 복근은 꾸민 것' http://news.joins.com/article/19006555 ¹³ 참고 | Falling in love with AI virtual assistants: a creepy love affair nearer than you think ¹⁴ 논문 | 내러티브 도구로서의 소품 활용 연구 - 영화 '그녀(her)'를 중심으로 ¹⁵ 참고 | http://movie.daum.net/movieDb/photoviewer?id=80976#950384 ¹⁶ 참고 | http://movie.daum.net/movieDb/photoviewer?id=44242#178388 ¹⁷ 참고 | RealDolls sex robots are ready to talk you into bed ¹⁸ 참고 | https://www.digitaltrends.com/cool-tech/realdoll-sex-robot-ai/ ¹⁹ 참고 | Relationships with Robots: Good or Bad for Humans? ²⁰ 참고 | The Good, The Bad and The Robot: Experts Are Trying to Make Machines Be "Moral" ²¹ 참고 | 테드 영상, alone together

카카오 AI 전문가 모집

Recruiting AI Experts

석·박사생을 위한 카카오 상시 연구지원 프로그램

Kakao I Research Supporting Program

모집분야	검색모델링	검색 질의 의도 파악, 문서 분류 및 검색어 관련성 연구 등
	검색통계	검색 로그 분석, 평가지표 개발, 사용자 행태 분석 등
	멀티미디어	딥러닝 기반 computer vision 및 음악 정보검색 알고리즘 연구/개발
	음성처리	음성 인식/합성 엔진 및 언어처리 개발
	자연어처리	형태소분석기, 스펠러, 맞춤법 검사기, NLU시스템, QA시스템, 대화형 에이전트 개발
	컨텍스트처리	기계 번역, 다국어 번역, 문서 자동 요약, 대화 처리 연구/개발
	추천기술	카카오 모든 서비스에 제공되는 추천 시스템 알고리즘 개발/분석
	추천기술(광고)	유저 이용 맥락에 최적화된 광고 추천을 위한 알고리즘 개발/분석
	AI 시스템 개발	이미지, 동영상, 음악 등 멀티미디어 관련 서비스 플랫폼 개발 등



모집대상	· 경력/신입 제한 없음
	· 학력/전공 제한 없음 (단, 각 모집 분야의 석사 이상에 준하는 전문 지식 요구)
	· 전문연구요원 전직 가능, 신규 편입은 별도 문의 (careers@kakaocorp.com)
모집시기	상시 영입
선발절차	서류 심사 → 온라인 테스트 → 오프라인 인터뷰(2회) → 최종 합격 (온라인 테스트는 모집 분야 별로 진행 횟수, 형태, 문제가 상이함)

모집대상	석사, 박사, 석박사 통합과정 재학생 및 입학 예정자
모집분야	AI 및 Computer Science-Engineering 관련 전 분야
모집시기	상시 영입
선발절차	서류 심사 → 온라인 테스트 → 오프라인 인터뷰(1~2회) → 최종 합격

프로그램	연구장학금	연구인턴십	해외 학회 참관비용 지원
			
	careers.kakao.com/ jobs/P-10461	careers.kakao.com/ jobs/P-10459	careers.kakao.com/ jobs/P-10476

지원공고	
최종 합격자 혜택·의무	<ul style="list-style-type: none"> · 학비/연구비 지원 (연 최대 2,000만원) · 수혜 중 학회 참관 비용 연 1회씩 지원 · 수혜 종료 후 카카오 정직원 의무 근무 조건(수혜 기간과 동일 기간) · 왕복항공권, 숙박비, 학회 등록비 전액 지원 · 참관 종료 후 2개월 내 카카오에서 참관 내용 공유 PT 진행 · 카카오 데이터 제공 · 근무지 부근 숙소 제공 (판교, 제주) · 인턴십 중 수행한 연구로 논문 게재 시 인센티브 최대 1,000만 원 추가 지급

마치며

AI와 기술의 결합은 어느덧 일상으로까지 확대됐습니다. 우리 사회 속에서 AI 외연이 확대되는 양상은 이번 호에서 다룬 미세먼지에 맞서기 위한 AI를 결합한 기술을 통해 또다시 확인할 수 있었습니다. 미세먼지는 집의 베란다 창문을 여닫거나, 가벼운 산책에 나설지 말지와 같은 소소한 일상의 선택에 직결이 되는 요인입니다. 미세먼지에 맞서는 기술로써 AI가 적용되는 현실은 'AI가 일상의 많은 부분에 이미 스며들어 있는 것은 아닌가?'라는 생각을 하게 합니다.

그렇다면, 일상에 AI가 스며들어 있는 깊이는 어느 정도나 될까? 이 호기심에 대한 질문을 찾기 위해, 일상과 AI를 주제로 한 외부 원고를 준비 중입니다. 일상의 여러 측면에서 적용되는 AI, 그 양상과 효과에 대한 이야기를 다음 호를 통해 보실 수 있을 겁니다.

봄이 오는 설렘에 비할 바는 못 되지만, 그래도 다음 호에 대한 기대감을 살며시 드렸으면 하는 바람에서 이번 호는 차기 호의 기획 아이템 하나를 소개드리며 조금은 다른 맺음 인사를 드립니다. 다음 호로 찾아뵙겠습니다. 감사합니다.

카카오 시리포트 원고 공모 REPORT

- 원고 마감일** | 매월 25일(익월호에 게재, 내부 판단에 따라 일정 변동 가능)
- 공모작 혜택** | 편당 50만 원
- 원고 주제** | 기술 전반 그리고 AI 산업 및 시장 분석
- 분량** | A4 5매 안팎(글자 크기 11pt, 행간 160%).
- 심사** | 카카오 시리포트 편집진과 카카오 AI 담당 기술 인력이 진행
- 투고처** | kakaoreport@kakaocorp.com
- 저작권** | 저작권법에 따른 정당한 범위 내에서 카카오 시리포트에 채택된 원고를 이용·가공·인용하시는 경우에도 반드시 출처를 밝혀주시기 바랍니다.

