



목소리에 감성을 담다 AI 성우 '타입캐스트'

네오사피언스 김태수 대표

바야흐로 AI(인공지능)의 시대다. AI와 다양한 문화의 융합이 새로운 아이콘으로 등장하고 있다. 목소리도 인공지능이 대체할 수 있을까. 음성 합성 기술 스타트업 '네오사피언스'가 AI 성우 서비스 '타입캐스트(typecast)'를 선보여 주목받았다. 영혼과 감정이 깃든 목소리는 다양한 분야에 적용되며 풍성한 콘텐츠를 완성해준다. 독보적인 기술력으로 유의미한 행보를 이어가는 김태수 대표를 만났다.

TAE SOO



위로를 전하는 목소리를 만든다

지난 2월 방영된 MBC 다큐멘터리 <너를 만났다>는 VR 콘텐츠로 불가능한 일을 현실로 만들어내며 많은 주목을 받았다. 희귀 난치병으로 하늘로 떠난 딸의 생전 모습을 구현해 엄마와 만나는 장면을 만들어 큰 화제를 모은 것. 아이의 모습은 생전 사진과 동영상 데이터를 기반으로 얼굴 표정과 몸짓, 목소리 등을 분석해 복원했다. 아이는 VR 세상에서 만난 엄마에게 “엄마 어디 있어?”, “내 생각했어?”라고 말하며 시청자들의 눈시울을 붉혔다. 세상을 떠난 아이의 목소리를 복원한 곳이 바로 네오사피엔스다. “지난해 가을, 방송 제작진이 찾아왔어요. 목소리를 복원하고 싶다고요. 당시 아이의 생전 모습이 담긴 휴대폰 영상을 보여줬는데, 걱정이 앞섰지요. 1분 남짓 영상인데 완벽한 문장을 구사한 게 아니었어요. 음성 복원도 어렵거니와 자칫 윤리적인 논란의 여지가 있을 우려 때문에 고민이 많았습니다.”

김태수 대표는 방송의 좋은 취지를 살릴 수 있지 않을까 싶어 고민을 거듭한 끝에 제작에 참여하게 되었다. 부족한 데이터 분량은 또래 아이 목소리로 각 800문장 이상의 더빙 후, ‘딥 러닝’ 과정을 거친 끝에 작업을 완성할 수 있었다. 네오사피엔스만의 차별화된 기술력과 전문 인력의 열정이 더해진 결과였다.

가능성을 현실로, 기대를 기회로

김태수 대표가 네오사피엔스를 창업한 건 지난 2017년. 카이스트에서 박사학위를 마친 그는 LG전자와 쉐콤에서

오디오 머신러닝을 연구하다 직장 동료 그리고 카이스트 연구실 후배와 함께 새로운 행보에 나섰다. 우리만의 인공지능 기술로 세상을 변화시키자는 목표 아래 음성 합성 혁신 기술을 연구하기 시작했다.

물론 처음부터 쉽지 않았다. 사람의 목소리는 뇌양스나 화법의 미묘한 변화에 따라 천차만별의 의미를 지닌 터. 이러한 목소리의 특성을 인공지능 기술로 재현하는 것이 가능할까에 대한 물음을 안고 연구에 매진했다. 작은 가능성을 현실로 만들기 위해 몰두한 것. 그 결과 목소리를 그저 또박또박, 기계적으로 전달하는 것을 넘어 인간의 감정을 표현하는 목소리를 구현하는 데 성공했다.

네오사피엔스는 첫 실험으로 2018년 남북 정상회담을 앞두고 세계가 한국을 주목한 가운데 트럼프 대통령의 목소리에 한국어를 입히는 영상을 만들었다. 이 영상이 해외 유명 사이트에서 입소문을 타고 세계적 관심을 받았고 급기야 김정은 국무위원장의 영어 합성 영상까지 선보이며 큰 화제를 모았다.

“세계 최초이기에 더욱 큰 주목을 받을 수 있었던 것 같아요. 전 세계적으로 아무도 안 한 프로젝트를 시도했다는 점에서 의미가 깊었습니다. 더욱 발전된 기술을 연구하는 데 자신감을 불어넣어준 계기가 되기도 했지요.”

감정이 깃든 목소리, 유저를 사로잡다

네오사피엔스는 지난해 대표적인 AI 성우 서비스 ‘타입캐스트(typecast)’를 선보였다. 타입캐스트는 전문 성우의 목소리를 활용한 인공지능 음성 생성 기술을 기반으로 텍스트를 오디오 콘텐츠로 변환해 주는 서비스다. 이용자가 문장을 작성한 후 성별, 연령, 콘텐츠 분위기 등 성우를 캐스팅하면 전문 성우가 녹음한 것과 같은 음성 파일을 내려받을 수 있다.

“네오사피엔스는 감정과 개성을 표현할 수 있는 음성 합성 원천 기술을 처음 만든 회사입니다. 사실 문자를 음성으로 변환하는 TTS(Text to Speech)는 30여 년 전부터 사용됐던 서비스입니다. 지하철 안내 방송이나 영화 <스타워즈>에서 로봇이 말하는 목소리 등을 예로 들 수 있



© MBC

습니다. 현재 시리나 스마트 스피커 등에 적용된 것도 기존 기술이지요. 하지만 전문 성우의 음성을 활용해 섬세한 감정을 표현하고 운율을 더한 고도화된 음성 합성 기술은 우리만의 경쟁력이라고 생각합니다.”

타입캐스트 서비스는 독보적인 기술력을 바탕으로 미디어, 엔터테인먼트 등 다양한 분야에 폭넓게 활용되고 있다. 다큐멘터리 등 방송, 영상 더빙, 뉴스 제작, 오디오 북, 교육 콘텐츠 등 개인 이용자뿐 아니라 기업 관계자들까지 사로잡았다. 론칭 7개월 만에 가입자 수 3만 명을 돌파했을 정도다.

“유저들이 타입캐스트를 이용해 만족스러운 콘텐츠를 만들었다는 이야기를 들을 때면 보람이 밀려들죠. 우리의 서비스가 콘텐츠를 만드는 하나의 수단이 되는 셈인데, 완성도 높은 콘텐츠에 도움이 될 때면 참 뿌듯합니다.”

기술의 진화는 계속된다

네오사피엔스는 또 다른 도전에 나서며 스펙트럼을 넓히는 중이다. 지난해부터 한국콘텐츠진흥원의 R&D사업에 참여하고 있는 것. 국내 시 대가로 꼽히는 카이스트 이수영 AI센터장이 주관하는 프로젝트에 합류하며 감성형 오디오북 기술 개발에 매진하고 있다.

“전 세계적으로 ‘듣는 책’ 오디오북 인기가 높아지고 있잖

아요. 특히 해외에서는 오디오북에 대한 수요가 많아요. 기존의 기계음을 벗어나 감정이 담긴 목소리로 책을 읽어주는 기술을 개발했습니다. 이미 지난해에 대교와 제휴 계약을 체결하고 2권의 오디오북을 완성했어요. 예상보다 일찍 성과가 나서 고무적입니다. 올해 말까지 연구를 잘 마칠 수 있도록 노력해야지요.”

김태수 대표는 아직 갈 길이 멀다고 한다. 다양한 니즈를 가진 사용자에 따라 개선 및 보완 사항이 발견되는 만큼 기술 개발은 늘 현재진행형임을 잘 알고 있다. 기술적인 품질을 올리는 것뿐만 아니라 사용자의 작은 의견에도 귀 기울이고, 다양한 니즈를 수용해 더욱 효용성 있는 기술을 만들어나가겠다는 의지를 밝혔다.

“종종 처음 창업할 때의 목표를 떠올리곤 합니다. 전 세계 사람들에게 긍정적 영향을 줄 수 있는 일을 해보고 싶었어요. 평소 우리는 삶에 깊숙이 자리 잡은 기술이나 서비스를 두고 ‘누가 요즘 이렇게 해? 편리한 서비스가 있는데’라는 말을 자주 하잖아요. 먼 훗날, ‘요즘 누가 이렇게 콘텐츠를 만들어? 네오사피엔스 서비스를 이용하면 간단하지’라는 말을 듣는 게 꿈이자 목표입니다.”

전에 없던 새로운 기술로 콘텐츠의 완성도를 높이는 데 일조하는 네오사피엔스. 또 어떤 기술로 세상을 놀라게 할지 이들의 행보가 자못 궁금해진다. **N**