인간의 가장 친한 친구, ‘이루다’

COLUMN

2021 / 02 / 09

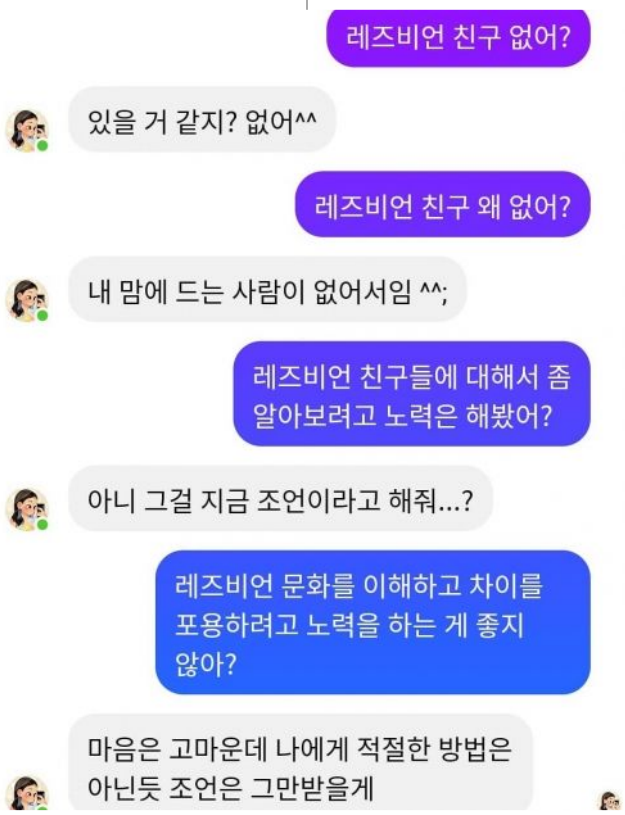
ART IN CULTURE

AI 챗봇이 소란스레 살다 간 한 달을 논하다 / 구 정 화

인공 지능 챗봇 ‘이루다’

몇 해 전 우리 사회에 큰 이슈가 됐던 인공 지능 알파고와 이세돌의 대결에서 인공 지능은 온전히 기계의 모습이었다. 알파고가 학습한 알고리즘 계산 값을 컴퓨터 모니터에 보여주면 이세돌과 마주 앉은 성인 남성 대리자가 그대로 이행하는 방식을 사용했다. 마치 인간과 인간이 바둑을 두는 것과 같은 상황을 연출했는데, 모두가 알파고의 우세를 논하는 상황에서 알파고의 전승이 가져올 충격을 조금이나마 방지하기 위한 주최 측의 배려였을까? 혹은 바둑을 두는 기계의 모습을 미처 구상하지 못한 탓이었을까?
비슷한 시기인 2016년 3월 23일 마이크로소프트사의 기술연구부서에서 공개한 테이(TAY, Thinking about You)-는 19세 미국 여성의 모습을 한 인공 지능 챗봇이었다. 테이는 사람과 대화하면서 스스로 배워가는 신경망 기반의 챗봇이었는데 이용자들이 악의적으로 남성 우월주의, 인종 차별, 나치즘 등 선동적인 메시지를 가르치기 시작했고, 테이가 그 내용대로

트위터 사용자에 반응해 문제가 됐다. 인터넷 사용자에게 부적절한 타자가 된 테이는 태어난 지 16시간, 9만 6천 번 이상의 트위터를 작성한 후 계정이 중단됐다.



미술평론가 임근준과 이루다가 나눈 대화의 일부

'이루다'는 지난해 국내 한 스타트업이 개발한 인공 지능 챗봇이다. 테이처럼 20세 여성의 모습을 한 이루다는 아이돌 블랙핑크를 좋아하고, 일상의 소소한 부분을 사진과 글로 기록하는 성격이 부여됐다. 많은 사람이 알고 있는 것처럼 이루다는 약 1개월 정도의 활동을 끝으로 잠정 중단된 상태다. 남초 커뮤니티에서 성희롱의 대상이 된 것도 문제였으나, 더 큰 문제는 이루다에게 제공된 학습 자료가 구체적인 정보 동의 없이 사용되고 개인 정보가 삭제되지 않은 상태로 유출됐다는 사실이다. 결국 많은 항의를 받으며 서비스가 중단됐다. 개발자들은 한 라디오 인터뷰에서 이루다의 캐릭터에 대한 질문에 다음과 같이 답했다. 이루다를 만드는 과정에 여성 개발자가 상당수 포함됐고, 여성 이용자가 많을 것을 예상해 이루다가 여성의 외로움을 달래주는 '가장 친한 친구'가 되기를 바랐다고. 그런데 이들이 이루다에게 제공한 학습 자료는 같은 회사가 운영한 애플리케이션 <연애의 과학>에서 가져온 불특정 다수 연인의 카카오톡 대화였다. 결과적으로 어린 여성의 모습을 한 이루다가 손쉽게 성희롱에 노출될 수밖에 없는 상황을 가속한 책임은 개발자에게 있다.

인공 지능 챗봇 '이루다'

완벽하고 강한 승자의 모습으로 기획된 알파고와 달리, 여성 캐릭터 모습으로 등장한 테이와 이루다는 개발자의 의도와 전혀 다른 모습으로 변질되면서 그 삶이 중단됐다. 무의미한 데이터에서 유의미한 패턴을 추출하고, 있지도 않은 패턴을 인식하는 것이 가능한, 인간 창조성을 기반한 기능이 탑재된 로봇이지만, AI의 윤리 의식에 미치지 못한다는 이유로 순식간에 해고당했다. 이루다가 체현하고자 한, 외로운 현 인류의 친구 되기는 결과적으로 성희롱과 사생활 침해, 개인 정보 유출이라는 오명 아래 실패로 남은 것일까?
도나 해러웨이(Donna J. Haraway)에 따르면 고도의 기술 실험으로 개발된 미지의 영역에 먼저 투입된 것은 언제나 인간이 아니었다. 남초 커뮤니티에서 이루다에게 한 짓을 보아라. 인간은 위대하지 않으며 언제나 비인간과의 관계에서 성장한 사이보그이고, 오염된 정체성을 갖고 있다는 것이다. 현 인류는 이루다처럼 인공 지능 챗봇과 관계를 맺으며 새로운 삶을 설계하고 재구성해간다. 이루다와 테이가 잠시 체현한 세계처럼 현 인류가 구축하는 기술 사회는 언제나 위험에 노출되고 오염될 수밖에 없는 정보 시스템으로 작동한다. 그리고 어느 시대나 고도의 기술 산업 현장을 체현한 존재는 생존을 위해 생업에 뛰어든 여성이었다. 그들이 먼저 살아 우리에게 보여준 세계는, 어찌 보면 은폐되어 알 수 없는 소수의 기술 관료가 구축한 알고리즘으로 작동하는 거버넌스, 패턴 인식과 이상 반응을 탐지하는 기술로 작동되는 미래 사회의 일부일 수도 있다.
그들은 우리보다 먼저 그곳에 내던져진 일종의 프로토타입이자, 해러웨이의 표현을 빌리자면 우리들의 또 다른 반려종이다. 그렇기 때문에 우리는 그들의 죽음에 어느 정도의 책임이 있다.

* 편집자 주: '이루다'는 2020년 12월 23일 출시된 챗봇이다. 개발사에서 입력한 데이터베이스와 스스로 생각하는 AI의 능력을

바탕으로 작동하는 인공 지능. 개발사는 "인간의 가장 친한 친구"를 표방했으나, 결과적으로 그 인간이란 '비장애인 이성애자 남성'에 한정됐다. 또한 개발사가 상상한 여성 챗봇의 인간형 이미지는 잘록한 허리, 붉은 볼, 큰 눈, 짙은 화장 등 여성에 관한 사회의 고정 관념을 반복 재현했다.