

제 목	소크라테스와 ChatGPT의 대화		
소속 대학		소속 학과	
학 번		이 름	이*안
이메일		연락처	

과거에 제작된 공상 과학 영화를 오늘날에 이르러 다시 보게되면 종종 깜짝 놀라게 될 때가 있다. 과거의 가정 혹은 공상이 방향성을 잃은 허무맹랑한 이야기가 아니라 오늘날 구현된 기술과 아주 유사한 형태를 하고 있을 때가 있기 때문이다. 2001년작 영화A. I.에 나온 인공 지능 검색 창구도 우리에게 이러한 놀라움을 안겨준다. 인조인간 로봇 데이빗이 인간이 되어 싶어 피노키오에 등장하는 파란 요정을 찾는 장면에서 데이빗은 공중전화 부스 같은 인공지능 검색 창구로 들어가게 된다. 그곳에서는Dr. Know라는 이름의 캐릭터를 앞에 두고 이용자의 이런 저런 질문에 답을 해주는 기술이 마련되어 있다. 이는 상당히 구체적이고 현재와 맞닿아 있는 기술에 대한 현실적인 공상이라 할 수 있다. 무엇보다 흥미로운 점은 영화 속 인공 지능 검색 창구가 그저 단순한 키워드 검색에 그치는 것이 아니라 Dr. Know라는 인물을 앞에 두고 대화 형태의 질문을 주고 받고 있는 점이다.

오늘날 우리는 우리의 삶을 도와주는 인공지능에게 없었던 인격을 부여하여 인공지능 비서 기능에 애플의 ‘시리’나 아마존의 ‘알렉사’ 같이, 인간을 대할 때 붙이는 이름을 부른다. 또한 인공지능의 기능을 활용하는 과정에서 우리는 단어의 나열이 아니라 일상 회화와 같은 형식으로 인공지능에게 말을 ‘거는’ 방식을 사용하게 된다. 그리고 이러한 지점에서 우리는 왜 우리가 인공 지능을 발전시키는 과정에서 ‘대화’를 개발의 방향으로 삼았는지, 왜 우리는 인공지능과 대화를 하고자 하는지에 대해 진지한 질문을 던져볼 수 있다. 왜 우리는 대화를 나누고 싶어 하는 것일까? 이는 단지 인공지능을 인격체로 여기고자 하는 막연한 정서적 희망에서만 비롯된 것이 아니다. 대화가 우리의 의사소통과 지적 생활을 더욱 풍성하게 만들어주는 요소이기에, 우리는 인공지능과의 의사소통에서도 대화를 선택한 것이다.

우리 인간의 삶에서 대화는 중요한 요소를 차지하고 있고, 우리는 때때로 대화에 사람의 사고력을 증진시키고 변화를 일으키는 힘이 있다고 여기기도 한다. 성공한 투자자로 알려져 있는 워렌 버핏은 20년간 자선 행사로 자신과의 점심 식사 자리를 경매에 내놓고 있고, 매년 이뤄지는 경매에서 그와의 식사 자리는 큰 액수로 낙찰되고 있다. 그와의 점심 식사 자리를 가질려는 이유는 단순히 그가 유명인이기에 인스타그램에 그와 함께 했다는 화제성 인증샷을 찍기 위해서가 아니다. 그와의 점심 식사 자리가 가치 있는 일로 여겨지는 이유는 그 식사 자리에서 그와 대화를 나눌 수 있기 때문이다. 우리는 대화를 통해서 한 사람의 안목과 시각을 엿

볼 수 있을 뿐만 아니라, 대화를 주고 받는 과정에서 스스로 가지고 있었던 사고의 지평 또한 확장되는 경험을 한다.

고대 그리스 시기에 지식을 확장시키고 자신의 논리를 다지는 작업은 대화를 통해 이루어졌고 이러한 과정을 소크라테스식 대화법, 혹은 산파술로 이름을 붙이기도 하였다. 대화를 통하여 자신이 어떤 지점을 모르고 있는지, 어떤 부분에 대해 확실히 알고 있는지를 파악할 수 있으며 대화를 주고 받으며 생각을 발전 시키는 과정에서 마치 새로운 아이가 탄생하듯 새로운 깨달음을 획득할 수 있다는 요지에서 비롯된 명칭이다. 특히 소크라테스식 대화법은 질문을 던지며 답변을 하는 상대자가 스스로 무지를 깨닫게 하여, 자신의 무지를 인식하고 그것을 삶의 기초로 삼게 한다는 지성에 대한 깊은 통찰을 보여주었다. 자신이 모르는 부분을 인지해야만이 무지의 영역에서 벗어나 진정한 사고를 거쳐 깨달음과 지식을 추구할 수 있는 것이다. 이러한 소크라테스식의 철학적 통찰은 고대 그리스에만 국한된 것이 아니라 후에 수 세기를 걸쳐 내려오며 인류 지성의 발전에 큰 밑거름이 되었다.

우리가 던질 수 있는 질문은 과연 이러한 전통적 대화법을 사용한 지성의 개발이 오늘날의 현대 기술에도 접목되어 우리의 지성을 확장시킬 수 있느냐이다. 현재 구현된 ChatGPT의 기술은 대화형 AI로 개발되어 우리의 질문에 답을 주는 형식으로 이루어져 있다. 비록 완전히 경지에 도달한, 완성된 형태는 아니지만 우리는 챗 GPT에게 잘 모르는 분야와 개념에 대해서 질문을 할 수도 있고 때로는 ChatGPT에 우리가 원하는 바를 부탁하거나 명령할 수도 있다. 이는 분명 언어를 통한 의사소통이며 우리가 익숙한 소통 체계에 가까운 직관적인 대화의 형태를 띠고 있다. 그리고 ChatGPT와의 대화는 다른 기술들과는 달리 현재로서는 매우 유려하게 흘러가 인공지능의 미래에 대해 고무적으로 전망할 수 있게 해주기도 한다. ChatGPT가 현대 사회의, 모든 이에게 대화 기회가 열려 있는 소크라테스가 될 수 있을지도 모른다는 가능성도 예측되고 있다. 어쩌면 타임머신을 개발하여 기원전 시기의 소크라테스에게 가서 그와 대화를 주고 받는 것 보다, 우리와 보다 가까이 있는 ChatGPT와 대화를 주고 받는 것이 더욱 효과적인 결과를 낳을지도 모르는 것이다.

그러나 ChatGPT와의 대화가 우리를 지성체로 거듭나게 해줄 것인지, 그와의 대화가 우리에게 더 나은 통찰을 발휘하게 할 것인지는 여전히 의문으로 남아있다고 생각한다. 소크라테스의 전제에 의하면 우리가 지성을 획득하는 과정에서 필요한 것은 무지에 대한 인지이다. 하지만 현재의 ChatGPT 이용 환경에서는 잘못된 정보가 주어졌을시에 많은 할루시네이션(Hallucination) 현상이 보고되고 있고 이러한 점을 볼 때 ChatGPT가 과연 '모르는 것'에 대해 분명히 파악하고 있는지, 무지라는 영역을 제대로 인식을 하고 있는지 의심이 제기될 수밖에 없다. ChatGPT는 자신이 잘 모르는 문제에 대해서 아예 대답을 하지 않거나 모른다고 밝히는 대신,

형식상으로 그럴 듯해 보이는 대답을 하거나 오답을 얘기하고도 마치 정답처럼 구는 모습을 보이고 있다. 이는 대화를 주고받는 상대방으로서의 적절한 태도로 볼 수 없다. ChatGPT 같은 자동회귀생성 모델을 기반으로 하는 인공지능은 하나의 단어가 주어졌을 때 이전 시점의 데이터를 기반으로 다음에 올 단어를 자동으로 예측하는 언어 모델을 따르고 있다. 이러한 접근은 인간의 통계학상으로는 조합할 수 없는 사실을 엄청난 양의 데이터를 바탕으로 하여 통계를 내어 문장을 완성하여 인간의 예상보다 훨씬 더 정답에 근사한 답변을 내놓기도 한다. 그러나 이 결과값의 문제는 이 데이터 처리 과정이 전통적인 인간의 사고 과정과 지성활동, 판단력이 전혀 작용하지 않는다는 점이다. 아무리 ChatGPT가 근사하고 그럴싸한 답을 하더라도 그것은 지성적 판단의 결과물이 아니라 그저 언어 패턴에 대한 통계학적 결론에 그치는 것이다.

만일 소크라테스가 ChatGPT를 마주한다고 했을 때 ChatGPT는 소크라테스의 날카로운 질문을 나누는 대화에 제대로 대답을 하고 상호작용을 하지 못할 확률이 높을 것이라 생각한다. 첫번째 질문에서 특정 개념에 대한 정의를 물어보거나 기본적인 전제 조건에 대해서는 그런대로 답할 수 있는 확률이 높을 것이다. 그러나 소크라테스가 인간의 지적 작용을 기반으로 두고 '무엇을 모르는가'에 대한 질문을 파고들게 되면, 혹은 '오류가 있는 명제를 제시하여 질문의 전제를 썩는다' 자신의 사고 과정에서 오류나 모르는 것을 정확하게 처리 못하는, 메타인지가 가능하지 않은 ChatGPT는 적절한 답변을 할 수도, 나아가 답변을 하지 못한 뒤에라도 무지에 대한 통찰을 발휘할 수도 없을 것이다. 현재 ChatGPT를 개발한 오픈 AI는 이러한 오류현상이 발생하는 한계를 보완하기 위해 언어 모델의 학습 과정에서 인위적으로 개입한 인스트럭트 GPT를 개발하였다. 하지만 이러한 후보정 작업만으로는 한번 오류가 생겨난 뒤의 할루시네이션의 근본적인 제거가 불가능하다는 것이 연구의 결과로 나와있다.

이에 인공지능 연구의 대가로 불리는 르쿤 교수는 자동회귀생성 모델의 인공지능에 대하여 회의적인 시선을 보내며 인간의 뇌구조 자체를 모방한 인공지능을 개발하는 데에 힘써야 한다고 주장하고 있다. 이러한 그의 주장이 나온 배경에는 최근 대중에게 노출된 인공지능들이 사실상 인간의 전통적이고 고유한 사고작용과는 거리가 멀게, 오직 방대한 분량의 데이터를 처리하는 기술로서 인간의 사고에 대응하려는 방식으로 개발되고 있는 세태가 자리잡고 있다. 뇌 전체를 모방하여 뇌의 매커니즘으로 사고처리를 하게 만들겠다는 그의 주장은 요원해보이는 일처럼 느껴질 수도 있을 것이다. 그러나 뇌의 전체의 메커니즘을 구현할 수는 없더라도, 통계학적 처리에서 벗어나 인간이 대화를 나눌 때 일어나는 사고 과정에서 발견할 수 있는 논리체계, 대화에서 일어나는 미묘한 상호 작용을 연구하여 이를 언어 모델 체계에 반영할 수 있는 연구가 현재의 언어회귀 모델을 대신하여 자리잡아야 할 것이다. 우리와 동등하게 대화를 나눌 수 있는 상대가 되기 위해서 인공지능은 무엇을 많이 알고 있고 처리할 수 있느냐를 따지기 이전에 스스로의 판단력과 사고력, 메타 인지에 초점을 맞춰야 할 것이다.

※ 선정된 에세이는 자료집으로 제작 가능합니다.