



## 인간과 AI, 그 신뢰성과 안전성

오승곤 TTA저널 편집위원장

인간의 지능(HI, human intelligence)과 인공지능(AI, artificial intelligence) 중 어느 것이 월등할까요? “미래 사회에는 강력한 AI를 가진 로봇이 인간을 지배할 수도 있다.”는 걱정이 현실화될 수 있을까요?

인간은 때때로 특정한 현상을 쉽게 일반화하고, 잘못된 기억을 사실인 양 표출하고, 편향된 사고와 편견을 가지고, 감정의 기복에 따라 이해할 수 없는 행동을 한다고 합니다. 뇌 연구에 저명한 한 교수는 이러한 인간의 행동을 볼 때 “인간의 뇌는 오류의 기계”라고도 했습니다. 반면 컴퓨터는 프로그램의 버그나 외부 바이러스 침입 등으로 오류를 일으키기도 하지만, 일 처리가 빠르고 계산도 정확하며 감정의 기복도 없습니다.

컴퓨터 기술이 빠르게 발전함에 따라 오늘날 ChatGPT 같은 AI가 일반화되고 있습니다. 또한 연구자들은 인간처럼 다양한 문제를 해결할 수 있는 능력을 갖춘 AGI(Artificial General Intelligence)를 개발하고 있으며, 나아가서는 인간의 두뇌와 능력을 훨씬 능가하는 수준에서 작동하는 ASI(Artificial superintelligence)의 출현도 회자되고 있습니다.

인류사회가 추구해 온 기술개발의 정점 중 하나는 ‘기계의 의인화(擬人化)’가 아닐까 합니다. 모든 기계가 인간처럼 보고, 듣고, 말함으로써 인간이 쉽게 일을 시키고, 인간이 시키는 일을 더 빠르고 정확하게 할 수 있도록 말입니다. 지금 세계가 기계에 AI를 심어 넣기 위해 노력하는 이

유도 여기에 있습니다.

역사적 경험으로 볼 때, 무엇이든 인간이 통제할 수 있는 범위에 있어야 그것으로부터 인간의 안전을 담보할 수 있습니다. 어떤 것이든 통제할 수 없으면 위험해지고, 이에 따라 사고와 재난이 발생합니다. 만약 미래에 ASI가 출현한다면 자연의 섭리에 따라 '세상의 최상위 포식자'라는 지위는 ASI가 이어받을지도 모를 일입니다. AI 기술이 하루가 다르게 발전하고, AI를 탑재한 더 똑똑한 로봇이 출현하면 할수록 AI 이용에 대한 걱정과 우려가 커지고 있습니다.

그러나 '앞으로 AI가 HI보다 월등해져서 인간을 지배할 것'이라는 막연한 우려 때문에 다양한 AI 활용과 AI 기술개발을 저해하는 것은 바람직하지 않습니다. 인간이 AI로부터 안전해지려면 오히려 AI에 대한 적극적인 연구가 필요합니다. 우리가 AI의 개발과 활용에 적극적으로 개입할수록 AI를 제대로 통제할 수 있습니다. AI의 신뢰성과 안정성은 제작자가 AI를 어떻게 만들고, 사용자가 AI를 어떻게 사용할 것인가에 달려 있다고 봅니다.

구체적으로 보면 이렇습니다. 제작된 AI가 당초에 의도한 기능과 동작을 제대로 하는지, 즉 품질의 정도가 신뢰성입니다. 이는 제작자의 책임하에 관리되어야 할 영역일 것입니다. AI를 제작 의도와 달리 사용자가 악용하거나 오용해서 위험과 사고가 발생할 소지가 없는지, 그 사용에 따른 위험도가 안전성입니다. 이는 사용자의 책임하에 통제되어야 할 영역일 것입니다.

최근 출시되는 자동차 대부분에는 복잡한 전자장비와 제어장치가 필수적으로 들어갑니다. 일종의 AI가 탑재된 셈입니다. 자동차에서 많이 논란이 되는 것 중의 하나가 '자동차 급발진' 문제인데, 우리는 수년 동안 일어나고 있는 이 급발진 문제가 신뢰성의 영역인지 안전성의 영역인지 아직 모릅니다. 즉 우리가 자동차의 AI를 완전히 통제하지 못하고 있기에 문제가 해결되지 않고 있는 셈입니다.

결국 고성능 컴퓨팅, AI 기능에 대한 신뢰성 검증은 인류의 안전과 지속가능성을 확보하기 위하여 첨단 기술을 바탕으로 매우 세밀하고, 다양한 형태의 꼼꼼한 체크가 수반되어야 할 것입니다.

TTA는 지난 1월 AI융합시험연구소 내에 "AI신뢰성센터"를 개소하여 AI 제품에 대한 신뢰성 검·인증 서비스를 본격적으로 제공하고 있습니다. 산업 전 분야에 AI의 융합이 급속히 확산되는 상황을 인지하고 우리 AI 제품의 품질과 경쟁력을 더 높일 수 있도록 관련 기업과 학계 등과 긴밀히 협력하여 시장 중심의 더 효율적인 신뢰성 검증 체계를 만들어 나가는 한편, AI 사용에 따른 안전성에 대한 프레임 개발과 정착을 위해서도 관련기관과 적극 협력해 나갈 것입니다.

이번 호 <TTA 저널>은 AI의 신뢰성을 주제로 다양한 전문가들과 함께 고민해 보려 합니다. 여기에서 펼쳐지는 논의가 우리가 만드는 다양한 AI제품의 품질을 더 높이고, 그 신뢰성과 안전성을 체계화해 나가는데 작으나마 보탬이 되었으면 합니다. 