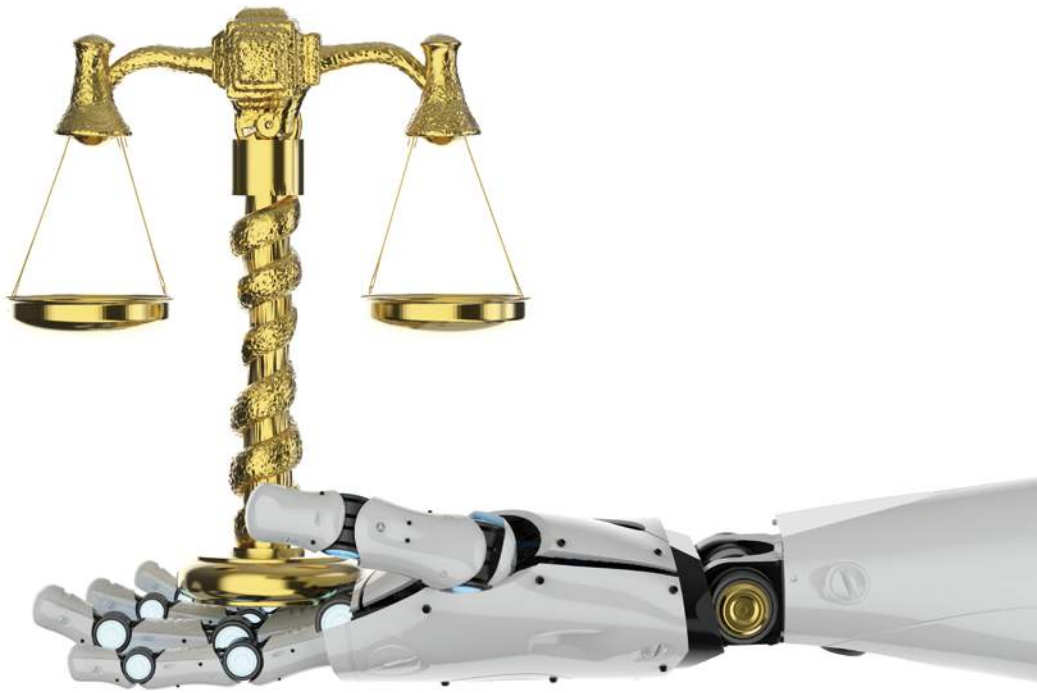


AI에 의한 차별, 무엇이 문제인가?

포털사이트의 야구경기 기사에 달려있는 댓글을 보면 사람들이 야구를 볼 때 심판의 스트라이크 존에 민감하게 반응한다는 걸 알 수 있다. 같은 공을 한번은 스트라이크, 다른 때는 볼로 판정하는 심판의 모습을 볼 때 마다 야구팬들은 차라리 인공지능(AI)이 심판을 맡았으면 좋겠다는 생각을 한다. 그런데 과연 AI가 심판을 하면 정확성, 형평성이 완전히 보장될까? 오히려 문제는 없을까?

글_황의관 선임연구원〈정책연구실 법제연구팀〉

인공지능이 인간을 차별한다?

AI와 관련한 부정적 문제 중 하나가 AI에 따른 차별 문제다. AI 기술 중 유력한 딥 러닝은 빅데이터를 활용한 통계적 분석을 기반으로 자기 학습을 통해 스스로 발전하는 것이다.

이세돌을 이겨 유명해진 알파고도 수많은 기보를 통해 학습하고 스스로 바둑을 두면서 그 결과를 다시 학습하는 방식으로 능력을 향상시킨 것이다. 따라서 당연한 이야기지만, 학습의 근간을 이루는 빅데이터가 편향성(bias)을 가질 경우 인공지능도 이를 공유할 가능성이 크다. AI를 사용한 차별이 발생할 수 있는 가능성이 상존하는 것이다. AI에 의한 차별은 여러 문제가 있지만, 이 글에서는 법적인 측면을 중심으로 살펴보도록 한다.

우리 헌법은 제11조에서 모든 국민의 평등권을 보장하고 있다. 하지만 여기서 평등은 절대적 평등이 아닌 상대적 평등으로서 '같은 것은 같게, 다른 것은 다르게' 취급하는 것을 의미한다. 따라서 단순히 AI가 차별적 처우를 했다

는 것만으로 그것이 평등권을 침해한 것이 되지는 않는다. 그리고 모든 국민이 가지는 기본권으로서 평등권은 본질적으로 대국가적 효력을 가진다.

즉, 차별의 주체가 공적단체(국가, 지방자치단체 등)일 경우 국민은 평등권을 침해한 국가 등의 행위에 대하여 기본권 침해를 주장할 수 있다. 평등권과 같은 자유권적 기본권은 국가 등에 대하여 방어적 성격을 가지기 때문에 국가 등의 침해행위가 있을 경우 이에 대한 회복을 청구할 수 있기 때문이다.

문제는 개인이 침해하는 경우다. 예를 들어, 사업자가 AI를 사용해 소비자를 차별하는 것이 소비자의 평등권 침해를 구성하는가? 소비자도 평등권을 가지지만, 사업자도 마찬가지다. 개인 사이의 문제는 원칙적으로 사적 자치에 맡겨져 있으므로 소비자는 사업자의 차별적 처우에 대해 평등권 침해를 주장하기 어렵다.

이러한 문제를 다루는 것이 기본권의 대사인적 효력인데, 현재 헌법학계와 헌법재판소는 이를 인정하고 있다. 따라서 AI를 사용한 사업자의 소비자에 대한 차별적 행위가 있는 경우, 평등권 침해를 구성할 수도 있게 된다. 물론, 개별 법률에서 사업자의 소비자에 대한 차별을 금지하는 규정이 있는 경우에는 헌법상 평등권 침해문제까지 논의될 여지는 없고, 해당 법률이 적용되면 그만이다.* 헌법상 평등권이 문제되는 것은 이런 법률이 없어 기본권을 직접 원용하는 경우를 가리킨다.

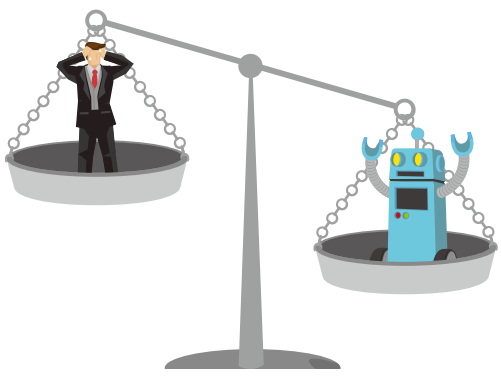
* 이러한 경우 헌법상 쟁점은 해당 법률이 헌법상의 원칙(비례의 원칙 등)을 준수한 합헌적 법률인지 여부이다.

AI의 잘못은 누구의 책임인가?

다음으로 중요한 법적인 쟁점은 누가 어떠한 방식으로 차별을 했고 누가 책임을 부담하는지에 대한 부분이다. 자동차를 운전하다 사고를 내면 일반적으로 운전자는 과실책임을 부담한다. 그럼 자율주행자동차가 사고를 내면 누가 책임을 부담하는가? AI를 사용한 차별의 문제도 비슷한 문제가 제기된다. AI와 사용자 중 누가 차별을 했고 누가 책임을 부담해야 하는가? AI에 의한 담합에서도 사업자들이 사용한 AI가 가격담합을 통해 소비자권익을 침해했다면, 누가 책임을 부담해야 하는가?

AI가 야기하는 부정적 결과와 그에 대한 책임의 문제가 여러 분야에서 다양하게 제기되고 있다. 이 쟁점은 차별적 행위의 주체가 누구인지의 문제와 차별적 결과에 대해 누가 책임을 질것인지의 문제이다.

전자는 AI를 사용한 행위가 사용자의 행위의 연장선으로 볼 수 있는지의 문제로 환원되고, 이는 AI의 역할과 이의 사용자의 역할 사이의 기여도 문제와 직결된다. 책임의 소재에 대한 문제는 행위의 주체문제와 연결되는데, 행위의 기여도에 따라 책임도 분배되어야 한다는 것이다. 예를 들어, AI가 3, 사용자가 2를 차별행위에 기여했다면, 책임도 그에 상응하게 부여해야 한다는 것이다.

다만, 여기서도 문제는 남는다. AI가 전부 기여한 경우 누가 책임을 질것인지의 문제와 위의 예에서 AI의 책임 부분인 3에 대한 손해를 어떻게 부담시킬 것인지가 문제다. 이에 대한 해결방안으로 몇 가지가 제기되고 있다. AI 사용자로서 전체의 책임을 부담해야 한다는 주장, AI의 사용에 따른 부담금을 징수하여 책임재산을 조성하자는 주장 등이 제기되고 있다.

AI가 대세로 훌쩍 다가온 오늘날 AI 활용에 따른 여러 장점과 함께 법적차원에서는 수많은 논의가 이어지고 있다. 이 글은 그 일부분을 그것도 질문을 제기하는 수준일 뿐이다. 차후에도 이런 문제에 대한 지속적인 관심과 연구가 이루어지기를 기대해 본다.