

인공지능(AI)의 인간화, 법적 인간화, 신뢰성: 인 간 중심의 관점과 전망

Personhood, Legal Personhood, and Trust- worthiness of AI: The Human-Centered Per- spective and Beyond

김건우 Gunoo Kim
광주과학기술원
GIST

국문요약

인공지능이 인간 삶에 파고들어 인간에게 미치는 영향이 더욱 커지면 커질수록, 인공지능의 도덕적, 법적 지위가 무엇인가 하는 주제는 하나의 근본적인 쟁점이 되고 있다. 그런데 많은 논자들은—법인격 개념 일반에 대해 그러했듯이—인공지능의 법인격에 대해서도 이를 인공지능의 인격성을 해명하는 문제로 다루어 왔으며, 이는 인공지능의 법인격성에 대한 찬반 진영 양쪽 모두 마찬가지였다. 그러나 필자가 보건대, 이러한 시도는 인격 개념과 법인격 개념 간의 일반적 관계에 관한 다음 몇 가지 사실과 논제를 통해 반박할 수 있다. (i) 인격 개념은 시대와 사상에 따라 형이상학적 인격 개념관과 경험과학적 인격 개념관을 포함한 다양한 견해를 포함하고 있기에 다중적이고 복합적이며 애매한 개념이며, 그 결과 많은 혼란을 초래해왔다. (ii) 법인격 개념에 관한 표준적 설명은 주로 인격 개념에 대한 (엄밀하지 않은) 전통적 이해에 의거하고 있다. (iii) (i)과

(ii)의 결과, 그러한 인격 개념에 내포된 중대한 다중성과 복합성, 그리고 애매성에 기인한 개념적 혼란이 고스란히 법인격 개념으로 전이되었다. (iv) 하지만 역사적으로나 개념적으로, 법인격 개념과 인격 개념 사이에는 의미심장한 간극이 있다. (iv) 따라서 법인격 개념을 설명하고자 할 때에 반드시 인격 개념에 의거할 필요는 없다.

이러한 논변으로부터 우리는 두 가지 결론적 함축을 이끌어낼 수 있다. 첫째, 인격 개념이 법인격 개념 일반에 대해 야기한 이같은 혼동과 신비를 걷어낸다면 “법적 비인간 인공 행위자”로서 인공지능의 법적 지위를 재고할 여지가 생긴다. 둘째, 인공지능의 법인격성을 이렇게 전향적으로 재검토할 때, 그것은 오늘날 공정성, 책무성, 투명성 등과 관련하여 아주 소중한 목표로 부각되고 있는 요청이라 할, 인공지능의 “신뢰성”을 확보하고 제고하기 위한 개념적 장치가 될 수 있다.

Abstract

As artificial intelligence pervades human life and its impact on human beings increases, the question of artificial intelligence’s moral and legal status becomes a fundamental

issue. Indeed, many commentators have discussed artificial intelligence's "legal personhood" as a question of clarifying its "personhood," as they did with the concept of legal personhood in general. This is the case with both the pros and cons of artificial intelligence's legal personhood. In my view, however, this line of discussion can be refuted through the following few facts and theses regarding the general relationship between the concepts of personhood and legal personhood. (i) The concept of personhood is a multiple, complex, and ambiguous one in that it contains a variety of conceptions of it, including the metaphysical conception and empirical conception, depending on the times and thoughts. Such multiplicity, complexity, and ambiguity have brought about much conceptual confusion. (ii) The shared account of the concept of legal personhood is mostly based on the rough traditional account of the concept of personhood. (iii) As consequences of (i) and (ii), the conceptual confusion as found in the concept of personhood has transferred to that of legal personhood. (iv) But historically and conceptually, there is a significant gap between the concepts of personhood and legal personhood. (iv) Hence, we do not necessarily have to account for the latter in terms of the former.

The above argument offers us two significant implications. First, if we remove the confusion and mystery caused by the concept of legal personhood in general, there is room for reconsidering the legal status of artificial intelligence as a "non-human artificial legal agent." Second, artificial intelligence's legal personhood reconsidered as above will arguably end up working as a conceptual device for securing and enhancing its "trustworthiness"—an invaluable goal that emerges today concerning its fairness, accountability, and transparency.

*Key word : artificial intelligence, bias, algorithm, fairness, discrimination, data