

인공지능, 권력변환, 세계정치 새로운 거버넌스의 모색

김상배

인공지능이 (국제)정치학의 관심을 끄는 이유는 새로운 ‘권력현상’의 출현을 엿보게 하기 때문이다. 사실 인공지능의 기술발달은 누가 어떻게 지배할 것인가의 문제를 되돌아보게 한다. 인공지능은 지배 메커니즘의 정교화를 의미하는 동시에 이에 대한 대항의 메커니즘을 구성하는 문제와도 연결된다. 이러한 와중에 인공지능 분야의 경쟁력을 확보하는 것이 국내외 정치에서 새로운 권력 게임의 핵심으로 부상하고 있다.

인공지능과 이를 둘러싼 권력 메커니즘을 (국제)정치학에서 주목하는 또 다른 이유는 이를 어떻게 다스릴 것인가, 즉 거버넌스의 문제 때문이다. 우선, 인공지능의 권력적 함의가 커지면서 이를 확보하기 위한 정책과 제도의 거버넌스를 정비하는 문제에서부터, 비대해지고 있는 인공지능 자체의 권력을 규제하는 제도와 규범을 만드는 문제가 관건이다.

이전에는 로봇이나 사이보그 정도로만 이해되던 인공지능이 최근 갑자기 우리 삶에 다가왔다. 기술공학 분야에서는 이미 오래전부터 진행되고 있던 현상이지만, (국제)정치학도 뒤늦게나마 인공지능에 대해 관심을 갖기 시작했다. 우리 삶에서 사이버 공간을 매개로 이루어지는 부분이 급속히 늘어나면서, 컴퓨터를 ‘언제 어디서나’ 사용하는 차원을 넘어서 컴퓨터가 인간 지능과 신체의 일부를 대신하는 과정이 진행되고 있다.

인공지능이 (국제)정치학의 관심을 끄는 이유는 새로운 ‘권력현상’과 연관되기 때문이다. 인공지능은 누가 어떻게 지배할 것인가와 밀접히 연결된 문제이다. 이러한 지배와 대항의 권력정치에 참여하는 주요 행위자들은 인간 행위자 중에 누군가일 가능성이 크다. 그런데, 인공지능의 게임에서는 그 누군가가 인간이 아닐 가능성이 있다. 이러한 인공

지능 게임의 특성은 (국제)정치학으로 하여금 이를 어떻게 다스릴 것인가, 즉 거버넌스의 문제를 고민케 한다.

국내적으로는 인공지능을 촉진 또는 규제하는 제도와 규범을 어떻게 만들 것이냐의 문제가 관건이다. 대외적으로도 미래 국가전략을 모색하는 차원에서 인공지능에 대응하는 정책을 마련하는 문제가 중요해지고 있다. 사실 선진국들이 인공지능을 주도하고 있는 현실 속에서 인공지능의 권력현상에 대비하는 것은 한국과 같은 중견국에게도 중요한 국가전략적 사안이다. 인공지능은 세계정치의 경쟁과 협력, 그리고 제도와 규범 형성의 문제이며, 더 나아가 세계정치의 구성원리와 작동방식의 변화와 연관되는 문제이기 때문이다. 이 글은, 이러한 문제의식을 바탕으로 인공지능이 야기하는 권력변환의 내용을 살펴보고 이에 대응하는 거버넌스의 모색 방향을 짚어보았다.

1. 들어가며

2016년 3월, 인공지능 알파고가 프로기사 이세돌과 벌인 바둑 대결에서 4 대 1로 승리한 사건은 인공지능에 대한 대중적 관심을 높이는 계기가 되었다. 이전에는 로봇이나 사이보그 정도로만 이해되던 인공지능이 어느새 급성장하여 결국에는 우리의 삶을 압도할지도 모른다는 우려마저 제기되었다. 컴퓨터나 소프트웨어 공학 분야에서 인공지능 기술은 이미 1950년대부터 시작되어 큰 부침을 겪어온 논란거리지만, (국제)정치학에서는 최근에 들어서야 뒤늦게나마 겨우 인공지능에 대

한 관심을 갖기 시작했다. 물론 기술발달이 미치는 영향에 대한 (국제)정치학계의 관심은 그 이전부터 시작되었다. 이 책의 저자들이 참여한 작업만 보더라도, 인터넷에서부터 시작해서 소셜 미디어에 이르기까지 기술발달이 미치는 영향에 대해 관심을 키워왔으며, 이러한 관심은 4차 산업혁명에 대한 논의로 접맥된 바 있다.(김상배, 2008; 김상배·황주성, 2014; 김상배, 2017)

최근 진행되고 있는 기술발달의 과정을 면면히 들여다보면, 인공지능이 우리 곁에 한층 다가와 있음을 절감하게 된다. 우리의 삶에서 사이버 공간을 매개로 이루어지는 부분이 급속히 늘어나는 현상과 병행하여, 컴퓨터를 ‘언제 어디서나’ 사용하는 차원을 넘어서 컴퓨터가 인간 지능의 일부를 대신하는 과정이 진행되고 있다. 더 나아가 컴퓨터는 책상 위의 데스크톱에서부터 무릎 위의 노트북을 거쳐서 인간 신체의 안으로 침투하고 있는 중이다. 최근 손쉽게 착용하는 웨어러블 컴퓨터를 개발하려는 시도가 한창 진행되면서, 미래의 컴퓨터는 아예 우리 신체의 일부로 이식될 것으로 예견되기도 한다. 결국에는 인간보다 지능이 우월한 컴퓨터가 출현하여 오히려 인간을 그 컴퓨터의 일부로 흡수하고, 그 인간(human)을 포스트 휴먼(post-human)으로 변신시킬지도 모른다는 전망마저도 나오고 있다.

이러한 일반적인 전망과는 달리, 인공지능이 (국제)정치학의 관심을 끄는 이유는 새로운 ‘권력현상’의 출현을 엿보게 하기 때문이다. 사실 인공지능의 기술발달은 누가 어떻게 지배할 것인가의 문제를 되돌아보게 한다. 인공지능은 지배 메커니즘의 정교화를 의미하는 동시에 이에 대한 대항의 메커니즘을 구성하는 문제와도 연결된다. 이러한 외중에 인공지능 분야의 경쟁력을 확보하는 것이 국내외 정치에서 새로운 권력게임의 핵심으로 부상하고 있다. (국제)정치에서 이러한 지배와 대

항의 권력정치를 주도하는 행위자들은 기존의 인간 행위자 중에 누군가일 것이다. 그런데 인공지능의 게임이 지니는 독특성은 그 누군가가 인간이 아닐 가능성도 있다는 데서 발견된다. 과연 이렇게 인간 및 비인간 행위자들이 참여하면서 벌어지는 인공지능 게임의 와중에 새로운 (국제)정치질서는 어떻게 출현할까?

인공지능과 이를 둘러싼 권력 메커니즘을 (국제)정치학에서 주목하는 또 다른 이유는 이를 어떻게 다스릴 것이냐, 즉 거버넌스의 문제 때문이다. 우선, 인공지능의 권력적 함의가 커지면서 이를 확보하기 위한 정책과 제도의 거버넌스를 정비하는 문제에서부터, 비대해지고 있는 인공지능 자체의 권력을 규제하는 제도와 규범을 만드는 문제가 관건이다. 이와 관련해서는 개별국가 차원에서 인공지능 경쟁에 효과적으로 임하기 위해 국내환경을 정비하는 문제 이외에도 이 분야의 국가 간 경쟁을 적절하게 규제할 국제 규범의 도입에도 관심이 집중되고 있다. 이런 점에서 보면 인공지능의 부상은 단순한 기술현상에만 그치는 것이 아니라 이를 둘러싼 국내외 정치의 '게임의 규칙'을 새로이 세우는 거버넌스의 문제라고 할 수 있다.

인공지능의 권력에 대한 기존의 논의는 다소 이분법적인 구도에서 단순하게 설정된 감이 없지 않다.(김대식, 2016) 한편으로, 인공지능 기술의 발달은 인간의 지능과 육체의 힘을 증대시켜 인간의 삶을 풍요롭게 하는 세상을 앞당길 도구로 그려진다. 인공지능은 인간의 지적 극한을 확장하고 인간의 수명을 연장하는 유토피아적 미래를 상징한다. 이에 비해 인공지능 기술의 발달은 기계와 인간이 대결하는 극단적 상황도 떠올리게 한다. 인공지능이 인간의 의도와 능력을 초월하여 그 창조자인 인간에 대해서 권력을 발휘하는 디스토피아를 초래할지도 모른다는 우려이다. 인공지능에 대한 대중적 인식을 반영하는 이러한 두 가지

시각은, 자칫 인공지능에 대한 지나친 기대감이나 막연한 불안감을 조장할 가능성이 있으며, 더 중요하게는 인공지능이 야기하는 권력변환을 이해하는 데 별다른 도움이 되지 않는다는 문제점이 있다.

이 글은 이러한 이분법적 구도를 넘어서 인공지능이 야기하는 권력변환의 내용을 따져보고, 이에 대응하는 거버넌스의 방향을 고민하였다. 이러한 문제의식을 바탕으로 이 글은 크게 세 가지 질문을 던졌다. 첫째, (국제)정치학의 시각에서 문제시되는 '인공지능'이란 무엇을 지칭하며, 이러한 인공지능을 어떻게 이해할 것인가? 둘째, 인공지능의 부상은 권력의 성격과 구조 변화에 어떠한 영향을 미치고 있으며, 이러한 변화에 어떻게 대응해야 할 것인가? 끝으로, 인공지능을 둘러싼 국가 간 경쟁과 협력의 양상은 어떻게 나타나고 있으며, 이러한 와중에 세계정치는 양적·질적으로 어떠한 변환을 겪고 있는가? 이 글은 이러한 질문에 대한 답을 찾는 과정에서 인공지능이 야기하는 권력변환과 거버넌스에 대한 탐구가 향후 (국제)정치학 연구의 핵심 아젠다가 되어야 함을 주장하였다.

2. 인공지능의 개념적 이해

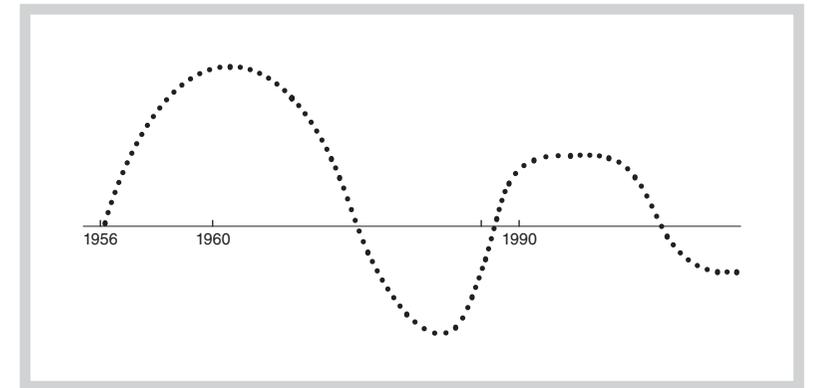
인공지능의 진화와 유형

(국제)정치학의 시각에서 볼 때, 인공지능으로 대변되는 현상이 중요한 이유는 인공지능 관련 기술이 이른바 미래 선

도 부문(leading sector)의 향배를 극명하게 보여주는 사례이기 때문이다. 인공지능이 주목을 끌기 이전에도 1980~90년대에는 반도체와 소프트웨어 기술이 이러한 관심의 대상이었으며, 2000년대에는 인터넷과 디지털 콘텐츠에 대한 관심이 부각되었고, 2010년대에는 소셜 미디어나 스마트폰과 같은 대상들이 (국제)정치학자들의 관심을 끌었다. 최근에는 선도 부문의 현황을 보여주는 화두로 클라우드 컴퓨팅, 사물인터넷, 빅데이터, 자율주행자동차, 드론, 로봇 등이 거론되고 있다.(로스, 2016) 그러던 와중에 알파고-이세돌 대국 이후에는 인공지능이라는 말에 초점이 맞춰졌으며, 최근에는 '4차 산업혁명'이라는 용어하에 이상의 모든 문제들이 집합되는 모양새를 보이고 있다.(하원규·최남희, 2015; Schwab, 2016; 김상배, 2017)

이 글은 이러한 상황 인식을 염두에 두면서도, 미래 선도 부문에서의 기술발달이 야기하는 권력변환과 이에 대응하는 거버넌스의 방향에 대한 논의를 이 책의 주제인 '인공지능'에 초점을 맞추어 펼쳐보자 한다. 이런 점에서 이 책에서 사용하는 인공지능이라는 말은 그 자체가 지닌 말뜻보다는 좀 더 포괄적인 외연과 내포를 지닌다고 할 수 있다. 역사적으로 좁은 의미에서 본 인공지능은 지금까지 두 차례의 붐을 겪었다. 1956년에서 1960년대가 제1차 붐이었다면, 1980년대가 제2차 붐이었다. 이러한 두 차례의 붐이 일었던 기간 인공지능이 가능케 할 장밋빛 미래에 대한 기대가 고조되었으며, 많은 기업들은 인공지능 연구에 뛰어들었고 막대한 정부예산이 투입되었다. 그러나 인공지능이 조만간 가능하리라는 기대는 쉽사리 실현되지 않았고, 그다음에 찾아온 세상의 반응은 더할 나위 없이 냉랭했다. 그러던 것이 2010년대에 들어 인공지능은 세 번째의 붐을 맞이하고 있다.([그림 1] 참조)

[그림 1] 인공지능 부침의 사이클



출처: 조성배, 「인공지능 기술의 원리, 현황, 전망」, 2015, p.3.

이러한 부침의 사이클에서 논의되는 인공지능의 범위와 차원은 매우 다양하지만, 이 글의 논의와 관련해서는 크게 '약한 인공지능(weak AI)'과 '강한 인공지능(strong AI)', 둘로 나누어 이해할 수 있다. 약한 인공지능은 도구를 활용해 온 연장선에서 인간의 인지능력이나 육체능력을 향상시키는 기술이다. 인간의 인지능력 전반을 수행하는 것이 아니라 구체적인 문제해결이나 추론기능을 수행하는 소프트웨어를 의미한다. 또는 디바이스의 신체이식이나 웨어러블 기술의 활용을 통해서 인간처럼(humanlike) 행동하는 기계를 만드는 기술이다. 한편 강한 인공지능은 단순한 도구를 넘어서 기계의 인지능력과 육체능력을 높여서 인간이 할 수 있는 지적 업무를 수행할 뿐만 아니라 이성적으로 사고하고 행동하는 시스템을 만드는 기술이다. 강한 인공지능에 대한 논의를 담는 가장 대표적 사례는 AGI(artificial general intelligence)이다. 인공지능 기술의 발전은 이러한 과정을 거쳐서 인간의 '의식'을 보완·대체하고, 인간의 '신체'를 보완·대체하는 단계로 발전하고, 더 나아가서 결국에는 '생명'을 불어넣는 단계로 진화할 것이 예견된다.

이렇게 다양한 양상으로 전개되는 인공지능 기술의 발달 과정에서 어느 부분을 강조하고 이를 어떻게 활용할 것이냐의 문제는 나라마다 다르게 나타난다. 다시 말해, 인공지능에 대한 정책적·사회적 담론은 각기 다르게 구성될 가능성이 크다. 예를 들어 이 책의 제2장은 알파고-이세돌 대국 이후 한국에서 출현한 인공지능 담론은 크게 셋으로 구분하여 소개하였다. 첫째는 ‘산업담론’으로 인공지능에서 새로운 산업의 기회와 새로운 사회구성을 기대하는 담론이다. 둘째는 ‘생존기술 담론’인데, 인공지능의 위력이 드세지는 환경에서 기계가 대체할 수 없는 인간만의 독특한 영역을 모색하겠다는 담론이다. 끝으로, ‘비관적 담론’은, 지능과 육체의 차원에서 기계가 인간을 대체 또는 능가하는 우울한 미래를 그리는 담론이다. 이러한 담론들에 의해서 구성되는 인공지능의 현실이 누구의 이익을 반영하는 방향으로 귀결될지를 분석하는 문제는 권력변환을 탐구하는 (국제)정치학의 중요한 논제가 아닐 수 없다.

인공지능을 보는 시각

인공지능을 어떻게 볼 것인가에 대한 논의는, 좀 더 근본적인 차원에서 인공지능을 보는 인식론과 존재론의 시각 차이에 따라서 크게 다르게 펼쳐진다. 이러한 이론적 시각의 차이는 인공지능의 현실적 발전에 영향을 미치는 기본 프레임을 규정한다는 점에서 주목할 필요가 있다. 이 글에서는 인공지능을 바라보는 시각을 크게 둘로 나누어 보았다. 이러한 구분은 인공지능 연구의 경향과도 맥을 같이 하는데, 인공지능을 ‘개체적인 행위자’로 볼 것이냐, 아니면 ‘환경적인 시스템’으로 볼 것이냐의 문제이다. 이러한 두 가지 시각 내에서도 인공

지능을 ‘도구적 변수’로 볼 것이냐, 아니면 ‘구성적 변수’로 볼 것이냐를 구분해서 보는 것도 의미가 크다. 이들 시각의 차이를 간략히 살펴보면 다음과 같다.

첫째, ‘개체적인 행위자’로 인공지능을 보는 시각이다. 이러한 시각에서 보면 인공지능은 기술발달로 인해서 생산된 새로운 산품이다. 자동장치(automata)와 같은 스마트 기기나 로봇을 떠올리면 이해하기 쉬울 것 같다. 이렇게 이해된 인공지능은 기술발달의 성과를 적극적으로 활용하여 인간 또는 기계의 능력을 향상시키는 차원에서 개발된 개체이다. 이렇게 인공지능을 개체로 보는 시각은 벤 피르첼(Ben Goertzel)의 AGI 접근에서 전형적으로 나타난다. 이는 인공지능 기술이 발달 속도가 폭발적으로 가속되어 인간의 지능을 훌쩍 뛰어넘는 순간, 즉 기술적 특이점(technological singularity)을 논한다. 그 시점으로는 대략 2050년경이 거론된다. 이러한 시각은 인공지능 기술의 발달로 탄생할 ‘초지능(superintelligence)’의 개체로서 포스트휴먼을 논한다. 즉, 인공지능의 발달로 인하여 인간의 생물학적 몸은 도태되더라도 첨단 기술에 의해 완전히 성능이 증강된 새로운 존재가 출현한다는 것이다.(Goertzel, 2015; Goertzel and Goertzel eds, 2015)

이러한 논의는 트랜스휴머니즘trans-humanism의 구성적 논의와 연결된다. 트랜스휴머니즘은, 인간 본성은 끊임없이 보다 나은 상태로 변화 가능하기 때문에, 인류는 기술발달을 적극적으로 활용하여 인간의 인지 및 육체 능력을 바람직한 방향으로 개선하여 인류의 무병장수를 지향해야 한다는 주장이다. 이러한 주장은 진화의 자연선택 메커니즘의 한계로 인간의 육체와 마음은 여러모로 불완전할 수밖에 없기에 생명공학이나 컴퓨터 공학기술을 활용하여 인류를 현재 상태보다 훨씬 더 진보된 ‘초인류’로 변화시켜야 한다는 사회운동으로 이어

진다.(이상욱, 2015, p.23) 이러한 트랜스휴머니즘의 사회운동은 옥스퍼드 대학의 철학자인 닉 보스트롬Nick Bostrom의 주도하에 1999년 ‘트랜스휴머니즘 세계연합’이 결성됨으로써 공식적으로 출범하였다.(Bostrom, 2014)

21세기 들어 트랜스휴머니즘은 그 세력을 확대하고 있다. 트랜스휴머니스트들은 저명한 과학자이거나 첨단기업의 경영진으로 활동하면서 관련 기술의 개발과 상용화를 이끌어가고 있다. 따라서 이들의 입장은 미래 기업의 비전을 제시하는 데에서 경험과 이론 측면에서 상당한 설득력을 발휘하고 있다. 또한 트랜스휴머니즘은 교육 영역의 시장화를 매개로 교육을 변혁하고 있으며, 정부의 교육정책에도 영향을 미치고 있다. 이와 관련된 가장 상징적인 사례로는 구글의 지원 아래 트랜스휴먼의 미래 비전을 선도하며 교육의 미래지향적 대변혁을 외치는 싱귤러리티 대학(Singularity University)을 들 수 있다. 이 대학의 총장은 바로 트랜스휴머니즘 분야의 거물인 레이 커즈와일Ray Kurzweil이다.(커즈와일, 2007)

이러한 과정에서 인공지능은 단순한 도구적 변수가 아니라 인간관계의 변환을 야기하는 구성적 변수이다. 이 책의 제6장은 인공지능의 도입이 야기하는 사람들 사이의 관계, 즉 인공지능을 매개로 하는 인간-인간 관계와 인간 정체성의 변화를 논했다. 여태까지 문제시된 인간의 사회적 관계가 인간-인간의 관계였다면, 새로운 개체적인 행위자로서 인공지능이 출현함에 따라 이러한 관계는 변화의 가능성을 맞고 있다. 인간-인간의 관계에 인공지능이라는 인간과 대단히 유사한 존재가 끼어들면서 생기는 변화이다. 제6장은 이렇게 인공지능을 매개로 한 인간-인간 관계 또는 인간-인공지능 관계의 변화를 가상현실(virtual reality) 기술의 사례를 통해서 살펴보았다. 이러한 변화는 다른 사람과의

거래적인 계약관계뿐만 아니라 감정적으로 소통하는 관계의 변화도 야기한다.

둘째, ‘환경적인 시스템’으로 인공지능을 보는 시각이다. 이러한 시각에서 보면, 인공지능은 빅데이터, 사물인터넷, 클라우드 컴퓨팅, 로봇, 바이오 기술, 나노기술 등이 융합되면서 출현하는 사이버 물리 시스템(CPS, cyber physical system) 또는 인간과 기술 사이의 커뮤니케이션 시스템으로 이해할 수 있다. 여기서 인공지능은 가상현실이나 시스템에 편재한 행위 주체로서의 알고리즘의 형태로 그 모습을 구현한다. 이러한 시각에 입각하면 인공지능을 단순한 개체적인 행위자라고 묻는 발상에는 문제가 있으며, 오히려 인공지능은 시스템에 배태된 프로그램으로 보아야 한다. 이렇게 이해한 인공지능은 행위자-네트워크 이론(ANT, actor-network theory)에서 말하는 행위자-네트워크(actor-network)인 존재라고 할 수 있다.(Latour, 2005)

이러한 시각은 프란시스 하일리겐Francis Heylighen이 제기한 글로벌 브레인Global Brain의 관념에서 나타난다.(Heylighen, 2015) 그에 의하면 초휴먼 지능은 ‘어느 하나의 인공지능 시스템’으로 귀결되는 것이 아니라, 인터넷으로 연결되어 모든 사람과 인공물을 가로질러서 분산되어 나타날 것이라고 주장한다.(Heylighen and Bollen, 1996) 글로벌 브레인이란, 지구상의 인간들에 의해서 구성된 컴퓨터와 지식정보, 그리고 이들을 연결하는 네트워크로 구성되는, 자기조직화의 논리를 따르는 ‘집합지성’의 네트워크를 은유적으로 부르는 말이다. 이러한 글로벌 브레인의 관념은 인터넷의 폭발적 성장으로 인해 현실화되고 있다.(Heylighen, 2011) 하일리겐의 글로벌 브레인은, 피르첼류의 AGI론에서 말하는 기술적 특이점과 대비되는 개념으로서, 분산적 특이점(distributed singularity)을 제시한다. 하일리겐에 의하면, 향후 미래의

기술 주도의 지능성장은 어느 한 장소나 개체에 집중되는 것이 아니라 지구 전체의 인간들과 인공물에 편재할 것이라고 한다.

네트워크 시대의 인공지능은 인간보다 뛰어난 지능을 갖고 인간의 신체와 통합됨으로써 항시 도처에서 사용되는 편재성(ubiquity)을 실현할 가능성이 있다. 이러한 과정에서 인간의 생물학적 몸은 도태되고, 첨단 기술에 의해 완전히 성능이 증강된 포스트휴먼으로서 ‘포스트 호모 사피엔스(post-homo sapiens)’라는 새로운 종의 출현이 예견된다.(하라리, 2015) 여태까지 인간은 선택의 여지없이 숙명적으로 신체를 통해 구현되었지만, 포스트휴먼은 사이버 공간에서 자기가 처할 여러 가상현실들을 자유롭게 선택할 수 있다. 이렇게 복수의 가상현실 속에 사는 포스트휴먼은 마치 현재의 개인용 소프트웨어가 하드웨어를 바꾸어도 작동하는 것처럼 인간과 같은 생체적 몸에만 의존하지 않고 여러 가지 물리적 기반을 바꾸어가며 삶을 지속한다. 이러한 과정에서 신체가 생물학적 수명을 다해 소멸한다 해도, 인간은 다른 컴퓨터로 자신의 삶을 업로드하여 영생하는 ‘신체 없는 종’이 된다는 것이다.(이종관, 2015, p.5)

사실 새로운 행위자 또는 편재하는 포스트휴먼 시스템으로서 인공지능의 출현은 다양한 방면에서 정치사회질서가 질적으로 변화할 가능성을 전망케 한다. 인공지능의 시대를 맞이하여 인간사회의 기본적인 구성원리나 작동방식의 변환에 대한 논의가 나오는 것은 바로 이러한 이유이다. 예를 들어 이 책의 제3장은 인지가본주의론의 관점에서 인공지능의 발달로 인한 자본주의의 변환을 살펴보았다. 구성적 변수로서 인공지능의 발달은 디지털 공유와 P2P 생산을 기반으로 하는 새로운 자본주의, 즉 ‘확장된 인지가본주의’의 출현을 부추긴다는 것이다. 이렇게 구성적 변수로서 인공지능을 보는 시각은 국제정치의 영역에도 적용된다. 이 책의 제9장은 인공지능의 발달이 인간중심주의에 기반을 두

고 있는 근대 국제정치의 구성원리와 작동방식 자체의 변화에 영향을 미칠 가능성을 논하고 있다.

3. 인공지능, 권력변환, 거버넌스

인공지능과 권력변환

인공지능의 발달과 도입은 여러 가지 측면에서 우리 삶에 변화를 야기하고 있다. (국제)정치학의 시각에서 인공지능이 미치는 영향을 탐구한 이 글이 특별히 주목한 주제는 권력변환이다. 인공지능 기술의 발달과 이로 인해서 발생하는 변화는 관련 행위자들에게 불균등하게 영향을 미치며 그 결과로 행위자들의 관계적 배열과 자원배분의 방식, 그리고 권력구조의 내용을 바꾼다. 여기서 주목할 것은 인공지능의 부상 과정에서 문제시되는 권력자원의 내용과 이를 둘러싸고 벌어지는 권력 게임의 성격이 종전과는 질적으로 상이하게 나타날 가능성이 있다는 점이다. 권력론의 관점에서 보았을 때, 인공지능과 관련하여 논하는 권력변환의 개념은 적어도 세 가지의 복합적인 의미를 가진다. 이 글은 주로 권력성격의 변화, 새로운 권력자의 부상, 권력구조의 변동의 세 가지 측면에 주목하였다.(김상배, 2010)

첫째, 인공지능의 부상은 (국제)정치적인 의미에서 관건이 되는 권력의 성격 변환을 야기한다. 기존의 (국제)정치권력과 비교해 보았을 때 인공지능과 관련하여 관심을 끄는 권력은 군사력이나 경제력 등과

같은 물질적 ‘자원 권력’의 시각을 넘어서 종류의 것이다. 인공지능 권력은 이를 개발하는 기술과 지식을 확보하는 능력과 관련되며, 여기서 더 나아가 인공지능의 작동 과정에서 활용되는 데이터를 다루는 능력에서 비롯된다. 다시 말해, 인공지능 권력은 데이터를 구성하는 ‘기술의 구조’에 해당하는 알고리즘을 설계하고, 이를 활용하여 ‘보이지 않는 데이터의 구조’, 즉 패턴을 읽어서 좀 더 교묘한 방식으로 감시하고 규율하는 지식권력을 의미한다. 이는 디지털 시대의 구조적 권력 또는 설계 권력이라고 개념화할 수 있을 것이다. 이러한 점에서 인공지능의 부상 은 (국제)정치권력의 구성요소와 작동방식을 변화시키고 있다.

최근에는 인공지능 권력 중에서도 알고리즘 권력 또는 코드 권력에 대한 관심이 늘어나고 있다. 이러한 인공지능 권력에 대한 논의는 알고리즘이 설계되는 과정에 작용하는 권력을 문제시한다. 알고리즘은 중립적으로 작동하는 것이 아니라 이용하는 사람 혹은 객체와의 상호작용 속에서 끊임없이 수정 및 조정되며, 알고리즘의 통제 논리는 이를 설계하는 사람에 따라 다르다. 따라서 알고리즘은 이데올로기의 산물로서 다분히 정치적으로 편향될 가능성을 지닌다. 이러한 시각에서 보면 가치중립적인 알고리즘이란 존재하기 어렵다. 제4장에서 상술하듯이, 알고리즘 편향성은 포털 사이트의 뉴스 노출 알고리즘, 콘텐츠 큐레이션과 필터 버블, 유색인종에 대한 불리한 예방적 치안, 재범 위험성을 평가하는 콤파스COMPAS, 우버 택시의 요금결정 알고리즘 등의 사례에서 나타난 바 있다. 제5장이 다루고 있는 인공지능에 의한 가짜뉴스의 생산 문제도 이러한 편향성을 보여주는 사례이다.

이러한 인공지능 권력이 작동하는 과정에서는 알고리즘이 규정하는 지배권력의 질서를 어떻게 볼 것이냐가 문제로 떠오른다. 이러한 알고리즘 권력은 중립이라는 이름으로 행사되지만, 이는 지배권력의 강화

및 재생산, 정교화 등에 봉사하는 경우가 많다. 이런 점에서 공적인 정치권력이 부당하게 인공지능을 장악하여 활용할 가능성에 대한 우려가 제기된다. 최근에는 시장권력이 인공지능의 힘을 활용하여 사적권력을 행사하는 것이 더욱 주목을 끌고 있다. 인공지능을 활용한 빅데이터 권력의 빅브라더화가 주목을 받는 것은 바로 이러한 맥락이다.(슈나이어, 2016) 이는 미셸 푸코Michel Foucault가 말하는 감시권력과 권력/지식, 거버넌털리티governmentality에 대한 논의와도 맥이 닿는다.(Foucault, 1980; 1991)

둘째, 인공지능이 권력변환에 미치는 영향은, 인공지능을 둘러싸고 다양한 행위자들이 경합하는 가운데 새로운 권력자가 부상하는 현상으로 나타난다. 인공지능의 신흥권력 게임에 민첩하게 적응하고 더 나아가 이러한 게임의 방향을 주도하는 행위자가 새로운 권력자가 되는 세상이 도래했다. 이러한 과정은 기존 권력자들 간의 ‘권력이동’으로 나타나기도 하지만 새로운 행위자가 권력을 얻는 ‘권력분산’의 형태로 나타나기도 한다. 예를 들어, 인공지능 분야의 권력게임은 전통적인 국가 행위자들이 벌이는 자원권력 게임의 양상을 넘어서, 주로 초국적 인터넷 기업들이 벌이는 신흥권력 게임으로 나타난다. 이러한 과정에서 (세계) 정치 신흥권력의 주체가 국가 행위자가 아닌 비국가 행위자들이 될 가능성이 거론된다. 사실 최근 인공지능 분야의 권력은 공적 영역이 아닌 사적 영역의 시장 권력 또는 자본 권력에 의해서 주도되고 있다.

이러한 권력분산의 과정은 인공지능이 야기하는 지식권력 관계의 재편에서 극명하게 드러난다. 일반적으로 전문가들은 자신들의 지식을 토대로 하여 위계적인 권력관계를 형성한다. 예를 들어, 전문가들은 자신들만 이해할 수 있는 ‘언어’를 통해 일반인들과의 관계에서 우위를 점해 왔다. 이를테면, 의사와 환자, 판사와 일반인의 관계에서처럼 말이

다. 그러나 인공지능이 우리 삶에 적극적으로 도입되는 과정에서 이러한 ‘전문적인 언어’들이 ‘일반적인 언어’로 치환될 가능성이 높아진다. 전문가와는 달리 인공지능은 쉬운 말로 처방하고 판결하기 때문이다. 이러한 현상이 지식의 대중화와 천박화를 불러오는 것처럼 보일 수도 있다. 그러나 겉으로 보기에 점점 쉬워 보이는 현상 뒤에 숨어 있는 알고리즘의 프로그램은 점점 더 어려워진다는 사실을 명심해야 한다.

더 나아가 인공지능이 이야기하는 권력분산은 인간이 아닌 행위자, 즉 비인간(non-human) 행위자로서 인공지능에 권력을 부여하는 현상을 이야기한다. 이는 앞서 살펴본 새로운 권력자로서 인간 행위자들 간의 경쟁의 문제를 넘어서, 비인간 행위자인 인공지능이 ‘주체’가 되어 역으로 인간을 지배할 위험성에 대한 논의를 의미한다. 인공지능이 인간들의 집합지성을 넘어서는 초지능(super intelligence)을 갖게 된다면, 그리고 인공지능이 인간의 의지를 넘어서 자의식을 갖게 된다면, 인공지능이 인간을 소외시키고 더 나아가 인간을 통제하는 세상이 도래할지도 모른다. 이러한 우려는 스스로 ‘프로그램을 설계하는 프로그램(programmable program)’의 출현에 의해서 현실화될 것이다. 어쩌면 탈인격화된 존재로서 인공지능과의 관계가 더 민주적으로 작동할 가능성도 있지만, 권력을 남용할 가능성이라는 점에서 인공지능이 사람보다 낫다는 기대 자체가 또 하나의 ‘신화’일 수도 있다.

끝으로, 인공지능의 힘을 등에 업은 새로운 권력자의 부상은 궁극적으로 기성 권력구조의 변동 가능성을 낳는다. 인공지능의 부상을 둘러싸고 벌어지는 권력게임은, 제6장에서 설명하듯이, 인간 사회의 질서를 변화시킨다. (국제)정치 차원에서 보면 인공지능의 부상은 정치권력의 구조 또는 세력분포의 변동을 야기할 가능성이 있다. 이러한 현상은 당연히 신규 세력 간의 갈등을 야기하는데, 휴고 드 가리스Hugo de Garis

는 이러한 갈등을 인공지능을 옹호하고 그 도입을 주도하는 ‘우주론자(cosmists)’와 인공지능의 도입에 저항하고 그 도입 과정에서 도태된 ‘지구론자(terrans)’의 갈등으로 묘사하고 있다. 이들의 갈등은 단순한 이해갈등의 차원을 넘어서 핵심적인 가치체계들 간의 갈등을 보여주는 점에서 흥미롭다.(Garis, 2015)

이러한 갈등은 선진국들이 벌이는 세계정치의 패권경쟁으로도 나타난다. 제8장에서 강조하듯이 인공지능으로 대변되는 선도 부문 경쟁의 승패는 패권의 향배에 영향을 미친다. 이런 점에서 최근 인공지능 분야에서 미국과 중국, 그리고 독일, 일본 등이 가세하여 벌이는 경쟁에 주목할 필요가 있다. 제7장에서 제시하였듯이, 인공지능 분야, 좀 더 넓게는 4차 산업혁명 분야에서 벌어지는 경쟁의 승패는 산업과 경제의 영역을 넘어서 군사·안보·외교 분야를 포함하는 미래 세계정치 구조 전반의 패권변동에 영향을 미칠 것이다. 인공지능과 4차 산업혁명 분야에 대응하는 미래 국가전략의 마련에 각국이 열을 올리는 것은 바로 이러한 이유 때문이다.(김상배 편, 2017) 요컨대, 인공지능을 둘러싼 경쟁은 세계정치에서 권력의 목표와 수단 및 성격을 변화시키고, 결과적으로 권력구조의 변환을 야기할 것으로 전망된다.

이러한 권력구조의 변동은 인공지능과 관련된 법제도의 형성 과정에도 반영될 것이다. 최근 인공지능 기술의 잠재력을 최대한 살리고 부정적 파급효과를 최소화하기 위한 법제도적 방안의 모색이 한창이다. 예를 들어 인공지능을 활용한 무인자동차의 법적 책임 문제가 논란거리이다. 무인자동차가 사고를 내면 제조사, 자동차 제작사, 관련 공학자, 차 판매자, 운전자, 차 소유자 중에서 누가 어떤 방식으로 법적 책임을 질 것인가? 그 법적 책임을 밝히는 문제는 관련 당사자들의 합리적 협의를 통해서 해결 가능할 수도 있지만, 인공지능과 무인자동차의 개

발 과정에 투영된 관련 행위자들의 이해관계와 이를 관철하는 권력의 구조적 현실이 반영될 가능성이 크다. 물론 이러한 과정에서 무인자동차가 특정 국가에서 활용되는 범위나 도입된 역사, 사람들의 인식 등이 작용할 것이다.(이원태, 2015)

인공지능과 거버넌스

새로운 권력현상으로서 인공지능의 부상은 이를 어떻게 다스릴 것이냐의 문제, 즉 거버넌스의 문제를 제기한다. 사실 인공지능과의 관계에서 거버넌스의 문제는 다각도로 이해할 수 있다. 일단 도구적으로 인공지능을 설정하고 이를 달성하기 위한 거버넌스 체계를 정비하거나 인공지능을 도구적으로 활용하여 지배 거버넌스를 강화하는 문제를 생각해 볼 수 있다. 그러나 현재 인공지능 거버넌스와 관련하여 관건이 되는 부분은 좀 더 구성적인 변수로 인공지능을 설정하고 인공지능이 거버넌스의 주체가 될 가능성을 경계하고 이를 방지하기 위해서는 무엇을 어떻게 할 것이냐의 문제로 집중되고 있다. 다시 말해, 인공지능에 대한 인간의 통제를 확보함으로써 더 나아가서는 인간과 인공지능이 함께 어우러지는 세상을 만들기 위해서 어떠한 종류의 거버넌스를 구축할 것인가가 관건이다.

최근 쟁점이 되는 것은 앞서 언급한 알고리즘 권력의 편향성을 규제하기 위한 거버넌스의 문제이다. 제4장에서 다루고 있는 바와 같이, 알고리즘이 지배하는 정치, 즉 알고크라시(algocracy)의 과도화에 대한 견제 차원에서 제기된다. 알고리즘 지배의 공정성과 알고리즘 규제의 문제, 특히 코드 권력을 규범적으로 통제하여 알고리즘의 책무성(accountability)을 묻는 문제이다. 인공지능 기술의 잠재력을 최대한 살리고 부

정적 측면을 최소화하기 위해 관련 법적 방안을 새로이 마련할 필요성이 발생한다. 그런데 인공지능 알고리즘의 책임 문제는 그리 간단하지는 않다. 개체적인 행위자로 인공지능을 간주하고 그 책임을 묻는다는 이분법적 발상 자체가 논란이 될 수도 있다. 인공지능의 알고리즘을 책임지는 주체로 설정할 수 있는지의 문제에서부터 인공지능의 코드 권력을 규범적으로 통제하는 것만이 대안인가에 이르기까지 다양한 문제가 제기된다.

그렇다면 점차로 더 지능적으로 되고 자율적으로 되는 기계, 즉 자율기계(autonomous machines)를 어떻게 규제할 것인가? 게다가 ‘개체적인 행위자’ 차원이 아니라 ‘환경적인 시스템’으로 이해되는 인공지능에 대해서도 책임을 물리고 규제하는 것이 가능할까? 여러 곳에 편재하면서 인간 수준의 판단과 추리 능력을 발휘하는 ‘버추얼 인간(virtual person, virson)’으로서 인공지능에게 어떻게 책임을 물릴 것인가? 이러한 버추얼 인간이 갖는 도덕적 책임 또는 기계윤리(machine ethics)는 무엇을 내용을 해야 하는가? 여하튼 기계가 점점 더 인간의 판단과 지능에 유사한 능력을 획득하게 됨에 따라서 기계에 대해서 자율성과 넓은 의미의 의사결정 권한을 주라는 압력이 늘어나고 있다.(Burke, 2015)

향후 기계에 권한을 부여하는 문제는 인간사회, 좀 더 정확하게는 인간-기계 사회에 좀 더 깊게 통합될 것이다. 이러한 과정에서 의사결정의 주체와 위임의 문제가 제기된다. 진화론의 맥락에서 보는 인간의 고유영역은 어디까지이며, 기계에게 끝내 위임하지 말아야 하는 영역은 어디인가? 인공지능에게 얼마만큼의 의사결정 능력과 권한을 위임하는 것이 적절할까? 연산능력과 처리속도, 알고리즘, 데이터의 패턴 읽기 등과 관련된 부분에서는 인간이 기계를 따라갈 수는 없다. 의료나 법조 업무처럼 ‘경계가 정해져 있는 지식을 다루는 분야’도 인공지능의 몫으

로 간주된다. 그렇다면 위임하지 말아야 하는 영역은 어디일까? 인간의 정체성을 위협하지 않는 정도로 기계에게 권한을 위임하는 문제가 향후 쟁점이 될 것이다.(이원태, 2015)

사실 인공지능의 사실상 코드 권력에 대한 시민권력의 통제가 최근 주목받고 있다. 이러한 사례로 ‘오프소스 네트워크 권력’으로서 블록체인의 기반의 거버테크(govern-tech)에 대한 논의에 주목할 필요가 있다. 최근 블록체인은 기업뿐만 아니라 정치조직에도 적용·확산되며 조직 변화를 이끌고 있다. 세계경제포럼(WEF)은 블록체인이 개별 및 국가 간 금융거래에서 사용하고 있는 보안 시스템보다 더 안전하고 투명하며 효과적인 거래방식이 될 수 있다고 역설하였다. 블록체인은 제3자의 중개가 필요 없는 P2P 네트워크를 가능케 하며, 경제, 금융뿐만 아니라 정치와 정부의 역할 변화 등에 영향을 미칠 것으로 예상된다. 이는 블록체인 기반의 분산자율조직(DAO, Decentralized Autonomous Organization)의 부상을 의미하는데, 인공지능의 ‘집중형 권력’을 통제하는 ‘분산형 권력’의 가능성을 전망케 한다.

이러한 논의는 인공지능과 민주주의에 대한 논의로 연결된다는 점에서 정치학의 관심거리이다.(이원태, 2015) 사실 인공지능의 부상은 다양한 측면에서 민주주의의 실현과 관련된 여러 가지 문제들을 제기한다. 이념으로서의 민주주의라는 시각에서 볼 때, 인공지능을 투표권을 가진 행위자로 인정하는 민주주의 이념이 등장할까? 이러한 와중에 인공지능을 둘러싼 또는 인공지능을 매개로 한 인간의 인간 사이의 이익 갈등은 인공지능 시대의 민주주의에 어떻게 반영될까? 그리고 궁극적으로 인공지능 민주주의 이념과 이익의 갈등을 제도화하는 과정을 누가 어떻게 주도하여 진행할 것인가? 향후 인공지능과 관련된 민주주의 비전은 기술 발달에 따른 비용 감소의 긍정적 측면과 알고리즘 권력의

일상화에 따른 지배의 정교화라는 위험 사이에서 모색될 것이며, 이 과정에서 ‘인공지능 정치’에 대한 민주주의적 성찰은 반드시 필요하다.

이러한 성찰의 과정에서 최근 강조되는 것이 ‘사전주의 원칙(pre-cautionary principle)’의 도입 문제이다. 특히 약한 인공지능의 실현과 관련하여 ‘사전주의 원칙’을 적용하는 거버넌스가 논의된다. 사전주의 원칙은 다양한 위험성이 제기되는 과학기술에 대해 충분한 사전 대책을 수립하고 연구개발 과정 중에는 끊임없는 모니터링을 통해 기술 발전 과정을 적절하게 관리해 나가야 한다는 입장을 대변한다. 인간능력의 강화를 목적으로 하는 약한 인공지능 분야는 이러한 사전주의 원칙이 잘 적용될 수 있는 사례이다. 예상되는 이익만큼이나 잘못 확산된 포스트 휴머니즘 기술이 가져올 부작용도 무척 크기 때문이다. 가까운 미래에 실현될 약한 인공지능 기술은 이러한 사전주의 원칙에 입각하여 선제적 정책과 끊임없는 모니터링을 시행하며 생산적으로 관리해 나갈 필요가 있다.(이상욱, 2015, p.24)

먼 미래에 실현될 강한 인공지능 기술에 대한 거버넌스 원칙을 단일하게 제시하기는 쉽지 않다. 예를 들어, 미래의 어떤 시점에 정말로 기계가 자의식을 갖거나 인간을 뛰어넘는 지적 존재가 되는 상황이 발생한다면 어떻게 대응해야 할까? 매우 발전한 인공지능이 인간과 평화롭게 공존할 것이라고 전망하기에는 너무 막연하다. 결국 미래에 특이점을 넘어설지도 모르는 인공지능이라는 존재가 얼마나 인간과 가치를 공유하는지가 중요하다. 일부 논자들이 지금부터라도 공감 능력이나 상호 신뢰, 복지 등과 같이 인간에게 중요한 개념을 인공지능에게도 주입하자고 제안하는 이유는 바로 여기에 있다. 궁극적으로 먼 미래에 실현될 강한 인공지능 기술과 관련해서는 인간이란 무엇인가, 그리고 인간적 가치는 왜 소중한가 등에 대한 보다 근본적인 성찰과 그에 근거한

거버넌스 체제의 모색이 필요하다.(이상욱, 2015, pp.24-25)

궁극적으로 인공지능 거버넌스를 위해서는 인간에 대한 근본적인 성찰이 병행되어야 한다. 이는 이른바 ‘네오휴머니즘(neo-humanism)’의 논의와 연결된다. 네오휴머니즘은 포스트휴먼의 도래를 향해가는 첨단기술의 거침없는 발전 과정에서 인간이 직면하고 있는 실존적·존재론적 허무를 극복하고 인간존재의 의미를 회복시키려는 철학적 움직임을 일컫는다. 근대 휴머니즘의 변종이라고 할 수 있는 네오휴머니즘은 인간의 계산적 도구적 합리성을 중심으로 한 근대 휴머니즘이 과학기술과 공모 관계를 형성하며 결국 인간 스스로를 포스트휴먼이라는 이름 아래 도태시키는 역설적 상황을 지적한다. 구체적으로 네오휴머니즘은 인간을 디지털화함으로써 영생의 포스트휴먼 미래로 향하는 트랜스휴머니즘에 반하여 인간의 고유한 가치를, 몸, 예술적 몰입, 나아가서 죽음 등에서 발견한다. 네오휴머니즘은 인공지능이 시뮬레이션할 수 없는 인간적 부분들을 긍정적으로 해석하는 철학적 작업을 수행한다.(이종관, 2015, p.7)

4. 인공지능의 국가전략과 세계정치

인공지능 분야의 국가 간 경쟁

인공지능 관련 기술과 산업의 중요성이 커지면서 국가 차원의 경쟁이 가속화하고 있다. 좁은 의미에서 본 인공지능뿐만 아니라 좀 더 넓은 의미에서 본 무인자동차, 사물인터넷, 드론산업, 빅

데이터, 클라우드 컴퓨팅 등을 둘러싼 경쟁은 해당 산업 분야에서 벌어지는 기업 간 경쟁인 동시에 국가 간 경쟁의 양상을 띠고 있다. 예를 들어, 제7장에서 살펴본 바와 같이, 미국, 영국, 러시아, 중국 등 군사 강대국들을 포함해 세계 주요국들이 인공지능을 접목한 무인전투기와 살상 및 정찰용 로봇 개발에 열을 올리고 있다. 미국은 2015년 무인 무기 시스템 개발에 53억 달러(약 6조 원)를 지출하였다. 러시아는 2020년까지 기관총과 감시 카메라, 센서를 장착한 로봇을 만들어 미사일 기지에 배치할 계획이다. 이스라엘은 민간 기업이 개발한 ‘자폭용 드론’을 구입하는 시도를 벌이고 있다고 알려졌다.

그러나 인공지능 기술은 군사 분야보다는 민간 분야에서 혁신을 주도하고 있다. 인공지능 분야의 기술경쟁을 이끌어 가는 행위자들은 전통적인 국가 행위자들이라기보다는 글로벌 다국적 기업들이다. 실제로 인공지능을 포함한 4차 산업혁명 분야의 선두주자들은 구글, 도요타, 소프트뱅크, 바이두 등과 같은 다국적 기업들이다. 예를 들어, 미국의 경우, 구글은 2001년부터 인공지능 및 관련 분야에 280억 달러(약 33조 원)를 투자하였다. 일본 도요타 역시 10억 달러로 인공지능연구소를 설립하였으며, 일본 소프트뱅크는 이미 인공지능 로봇을 상용화해서 백화점 등에 배치한 상태이다. 중국에서는 인터넷 포털 회사인 바이두가 3억 달러를 투자해서 미국 실리콘밸리에 연구소를 설립하기도 했다.

이렇게 진행되고 있는 인공지능 분야 경쟁의 면면을 자세히 들여다 보면, 이는 단순한 제품경쟁이나 기술경쟁이 아니라 플랫폼 경쟁의 양상을 드러내고 있음에 유의해야 한다. 인공지능 분야의 경쟁에서는 하드웨어보다는 소프트웨어와 알고리즘을 장악하는 것이 핵심이 된다. 다시 말해, 경쟁자보다 더 우월한 알고리즘을 보유한 업체가 지속적으로 경쟁우위를 유지할 수 있고, 그 과정에서 경쟁우위 요소인 데이터를 더 많이 확

보하여야 다시 더 많은 플랫폼 참여자를 확보하는 것이 가능하다. 이러한 게임의 기본적 성격은 물량과 품질의 게임이 아닌 그야말로 승자가 독식하는 '규모'의 권력 게임으로 발전할 가능성이 매우 크다. 그 결과 네트워크 효과에 따르는 일정 수준의 시장 지배력을 가진 몇몇 플랫폼 간의 경쟁이 각 산업별로 일반화할 가능성이 있다.(김상배, 2017)

인공지능 분야에서 발생하는 이러한 경쟁양식의 변환은 4차 산업혁명 시대의 국가전략 전반에도 투영될 가능성이 있다. 특히 인공지능 분야 기술역량 개발과 인재양성 지원을 위해 효과적인 국내 정책과 제도를 마련하기 위한 경쟁이 벌어진다. 이는 앞서 언급한 인공지능을 위한 국내 거버넌스의 정비 문제와도 연결되는데, 이 분야의 첨단기술 개발을 지원하는 효과적인 정책을 마련하고 제도를 구축하는 것이 관건이 된다. 이는 진화생물학에서 말하는 적합력(fitness)의 관점에서 변화하는 대외 환경에 적응하기 위한 각국이 모색하는 생존전략의 차원에서 이해할 문제이다. 구체적으로 이는 기술환경의 변화에 적응하여 살아남기 위해서 국내 기술혁신체제를 정비하는 문제로 나타난다.

이렇게 정책과 제도, 혁신체제를 구축하는 문제는 인공지능이나 포스트 휴먼 기술의 현재와 미래를 둘러싼 국가별 담론의 차이에서도 발견된다. 예를 들어, 미국과 일본, 독일이 강조하는 인공지능 담론은 다소 다르다. 미국의 담론은 클라우드 환경에서 전 세계의 공장이나 제품에 대한 데이터를 수집하고 중앙의 서버에 데이터를 축적하며 인공지능으로 처리하는 모델이다. 앞서 구분한 인공지능을 보는 시각에 의하면, '환경적인 시스템'으로 인공지능을 보는 담론에 가깝다. 이에 비해 일본은 자동차 산업 등의 강점을 바탕으로 로봇 기반 인간접점 시장 공략 전략을 취하고 있으며, 이러한 과정에서 일본이 강조하는 것은 로봇이다. 이는 '개체적인 행위자'로서 인공지능을 보는 모델에 가깝다. 독

일이 지향하는 인공지능 담론은 그 중간쯤에 위치한다. 독일의 인더스트리 4.0 모델은 제조 현장의 생산 설비와 로봇 등 현실 세계의 강점을 지렛대로 현장 데이터의 네트워크화를 통해 새로운 사이버-물리 시스템의 구축을 지향한다. '환경적인 시스템'과 '개체적인 행위자'의 복합 모델인 셈이다.(김상배, 2017)

인공지능 분야의 경쟁은 커뮤니케이션 분야에서도 나타난다. 제5장에서 살펴보듯이, 인공지능을 활용한 기사 작성이 국내정치뿐만 아니라 국제정치에서도 논란이 되고 있다. 인공지능에 의한 신문기사 작성은 기상 및 재난정보와 같은 속보 대응, 단순한 수치와 통계를 제공하는 경기 결과 보도 분야에 활용되고 있다. 언론은 최신 뉴스를 보여주거나 간단한 질문에 답하는 형식의 추천 뉴스를 제시한다. 인기 있는 기사 리스트와 독자들의 호응도를 분류해 콘텐츠를 생산하는 등 효율성을 추구하며, 이는 대민관계나 공공외교에 활용된다. 그런데 이러한 과정의 부정적 측면도 만만치 않다. 의도적 거짓 댓글 등을 대규모로 생성하는 프로그램을 이용한 심리전이 논란거리이다. 기사작성 로봇을 통한 타국 선거에도 개입하여 논란이 되고 있는데, 러시아는 사이버 공간에서 가짜뉴스를 이용해 미국 대선, 프랑스 대선, 독일 총선, 스페인 카탈루냐 독립투표 등에 개입한 것으로 알려졌다.

인공지능과

세계정치의 변환

인공지능의 발달은 글로벌 생산 네트워크의 변환에도 영향을 미치고 있다. 제8장에서 다루었듯이, 인공지능이 도입되면서 제조업 분야를 중심으로 리쇼어링reshoring 현상이 발생한다. 인간이 하는 일의 대부분이 기계를 기반으로 한 기술혁신으로 대체되면서

제품 원가에서 인건비가 차지하는 비중이 미미해질 가능성이 있다. 특정 제품을 제외하곤 자국 내에서도 얼마든지 저렴한 비용으로 생산이 가능하므로 국제분업이 줄어들고, 그 결과 무역량이 급속히 감소하게 된다. 이러한 전망은 주로 '개체적인 행위자'로 인공지능을 보는 시각에 기반을 둔다. 그러나 '환경적인 시스템'으로 인공지능을 보면, 제조업 제품의 초국적 유통이 감소하는 현상보다는, 데이터와 정보의 초국적 유통은 오히려 증가할 가능성이 있다. 빅데이터나 클라우드 컴퓨팅, 인공지능 등을 활용하여 초국적으로 활동하는 다국적 기업들의 비즈니스는 더욱 늘어날 것으로 예견된다. 그렇다면 이러한 과정에서 발생하는 개인정보의 초국적 흐름과 빅데이터의 초국적 활용, 클라우드 컴퓨팅의 초국적 제공 등의 이슈는 어떻게 다룰 것인가?

최근 이 분야의 초국적 활동을 규제하는 국제규범 형성의 문제가 쟁점으로 부상하고 있는데, 이와 관련하여 국가별 입장의 차이에 주목할 필요가 있다. 미국은 자본의 시각에서 초국경 정보의 자유로운 흐름을 옹호하는데, 규제를 최소화하고 이에 맞는 세계질서를 만들려는 글로벌 패권국의 의도가 반영된다. 아주 민감한 분야를 제외하고는 국경 간 이동을 자유롭게 하자는 것이다. 중국, 러시아, 개도국들은 국가의 권리라는 시각에서 접근한다. 이들 국가는 선진국 기업들의 침투에 의한 자국 '정보주권'의 잠식을 우려한다. 인공지능 기술역량과 분석능력을 보유하고 있는 선진국들이 이를 부당하게 활용하여 권력을 행사할 가능성에 대한 우려가 제기된다. 유럽의 경우에는 페이스북을 상대로 한 맥스 슈렘스의 소송과 셰이프하버 협정의 무효화 사례를 볼 수 있듯이, 단순히 국가주권의 강화라는 차원을 넘어서 개인의 권리 또는 시민주권을 강조한다.

이러한 각국의 입장 차이는 개인정보 보호에 대한 각국의 이념적 ·

제도적 차이를 바탕으로 깔고 있다. 각국마다 개인정보 주체의 동의권 행사방식에 대한 법해석과 제도 운영이 다르다. 좀 더 근본적으로는 빅데이터의 중요성에 대한 각국의 인식과 제도, 언론의 자유와 개인정보의 우선순위 등에도 큰 차이가 있다. 미국이 언론의 자유를 더 중요시하는 반면, 유럽에서는 오히려 개인정보가 더 중요하다. 중국은 좀 더 큰 차이를 보이는데, 국가안보를 이유로 정부의 '동의 없는' 개인정보 수집 및 감청이 정당화된다. 빅데이터 국제규범에 대한 논의가 진행되면서 빅데이터 관련 규제정책의 도입과 정책 및 관행의 표준화 필요성이 제기되는 것과 동시에 각국의 제도와 문화에 내재한 이러한 차이들이 수면 위로 떠오르고 있다.

한편 최근 군사 분야의 인공지능 기술발달은 기계가 나서서 인간을 공격하고 전쟁을 벌이는 가능성에 대한 우려를 낳고 있다. 인공지능을 접목한 로봇은 폭탄을 투하하거나, 센서와 위치추적장치(GPS)를 이용해 적군의 동태를 정찰하는 기능을 수행하는 등 다양한 목적으로 활용될 수 있다. 그러나 자칫하면 스스로 결정해 살상을 저지르는 '킬러 로봇'이 될 수도 있다는 우려가 높아진다. 이러한 맥락에서 인공지능이 인간에게 해를 끼치지 못하도록 국제규범을 마련하자는 목소리가 높아지고 있다. 이는 인공지능 기반 자율무기에 대한 규제의 문제로서 킬러 로봇 금지 운동, 윤리적, 법적 이슈 등이 제기된다. 좀 더 관심의 대상이 되는 것은 인공지능을 활용하여 초국적으로 활동하는 다국적 기업들의 활동에 대한 규제인데, 이는 앞서 언급한 빅데이터와 개인정보의 활용, 클라우드 컴퓨팅 등의 문제와 연관된다.

인공지능이 세계정치의 변환에 미치는 영향은 근대 국제정치의 틀 안에서 벌어지는 국가 간 경쟁의 가속화나 새로운 규범의 형성 문제를 넘어설 가능성이 있다.(Cudworth and Hobden, 2011) 제9장에서 살

해보듯이, 인공지능의 부상 은 근대 국제정치에서 전제로 하고 있는 계몽주의와 인간중심주의, 즉 근대 국제정치의 합리성의 전제를 넘어서는 변화의 지평을 열 가능성이 있다. 인공지능의 시대를 맞이하여 인간중심주의 틀로는 설명할 수도 없고 통제할 수도 없는 현상이 발생한다. 특히 인공지능, 특히 포스트휴먼의 부상에 대한 논의는 계몽주의 기획 속에 숨어 있는 인간중심주의를 넘어서는 새로운 지평에서 이해할 필요가 있다. 예를 들어, 이른바 자율무기체계(AWS, autonomous weapon system)의 발달은 미래전쟁의 변화를 야기하며 이는 인간 중심의 전쟁 수행방식의 변화도 발생시킨다. 사이버 공간의 봇넷 공격, 악성코드, 무인기, 로봇/사이보그 등에 대한 논의는 인간이 아닌 행위자, 즉 비인간 행위자들이 행위능력을 갖고 세계정치의 전면에 부상하는 스토리를 창출한다.

5. 맺으며

이 글은 최근 기술발달의 화두로 관심을 끌고 있는 인공지능이 미치는 (국제)정치적 영향을 권력변환에 초점을 맞추어 살펴보고 그 연장선상에서 인공지능에 대한 새로운 거버넌스의 필요성을 지적하였다. 현재 거론되고 있는 인공지능은 인간의 인지 및 육체능력을 향상시키는 도구에서부터 인간처럼 사고하고 행동하거나 더 나아가 인간과 같은 존재를 만드는 기술에 이르기까지 다양한 형태로 발달을 거듭하고 있다. 이러한 인공지능은 인간의 의식과 신체를 보완·대체하는 단계

를 넘어서 장치 생명을 붙여넣는 단계로 진화할 것이 예견된다. 결국에는 인간보다 지능적인 컴퓨터가 출현하여 오히려 인간을 그 컴퓨터의 일부로 흡수하고, 인간이라는 종을 포스트 휴먼이라는 '신체 없는 종'으로 변신시킬지도 모른다는 전망마저도 나온다.

이 글은 이렇게 발전하고 있는 인공지능을 보는 시각을 크게 둘로 나누어 보았다. 먼저, '개체적인 행위자'로 보면, 인공지능은 기술발달을 적극적으로 활용하여 주로 인간 또는 기계의 능력을 향상시키는 차원에서 개발된 개체이다. 이러한 시각은 기술이 발전을 거듭하여 인간의 지능을 훌쩍 뛰어넘는 순간, 즉 기술적 특이점의 도래 가능성에 초점을 둔다. 한편 '환경적인 시스템'으로 보면, 인공지능은 빅데이터, 사물인터넷, 클라우드 컴퓨팅, 로봇, 바이오 기술, 나노기술 등의 기술이 제조업 및 커뮤니케이션 네트워크와 융합되면서 출현하는 시스템이다. 이러한 시각에서 보면 인공지능은 '어느 하나의 개체'로 집약되는 것이 아니라, 인터넷으로 연결되어 있는 인간과 기술이 분산적으로 만들어 내는 '글로벌 브레인'의 복합 네트워크이다.

이러한 인공지능이 (국제)정치학의 관심을 끄는 이유는 권력의 성격과 행위자 및 구조의 변환에 영향을 미치기 때문이다. 사실 미래의 새로운 기술 패러다임을 상징하는 인공지능의 발달은 누가 어떻게 지배할 것인가의 문제를 묻게 한다. 인공지능은 지배의 메커니즘을 정교화하는 문제인 동시에 이에 대항하는 메커니즘을 창출하는 문제와도 연결된다. 이러한 지배와 대항의 권력정치에 관여하는 주요 행위자들은 인간 행위자 중에 누군가일 가능성이 크다. 그런데 인공지능의 게임에서는 그 누군가가 인간이 아닐 가능성이 있다는 점 때문에 권력변환의 새로운 차원을 고민하게 된다. 최근 인공지능 기술권력의 비대화를 규제하고 새로운 거버넌스를 모색하려는 고민이 깊어지는 것은 바로 이

러한 이유 때문이다.

인공지능의 권력적 함의가 커지면서 이 분야의 주도권을 놓고 벌이는 국가 간 경쟁이 심화되고 있다. 인공지능의 발달이 세계정치 권력의 성격을 변화시키고 이를 둘러싼 경쟁의 외중에 권력구조가 변화할 전망마저도 낳고 있다. 무엇보다도 인공지능의 알고리즘 기술개발과 시스템 확보를 뒷받침하는 역량의 구비가 새로운 국가 경쟁력의 요소로 부각되고 있다. (국제)정치학의 시각에서 볼 때, 이 분야의 경쟁은 해당 기술과 산업 분야에서 벌어지는 기업 간 경쟁인 동시에 좀 더 넓은 의미에서 본 국가 간 경쟁으로 인식되어야 한다. 더 나아가 인공지능의 발달은 경쟁의 양식 변화나 권력구조 변동의 차원을 넘어서 미래 세계 정치의 주체와 구성원리 및 작동방식의 질적 변화까지도 예견케 한다.

한국도 인공지능 분야의 발달과 여기서 파생되는 국내외 질서 변동에 대응하여 국가전략 차원의 육성·지원의 정책을 마련해야 할 것이다. 인공지능으로 대변되는 신흥 선도 부문에서의 승패는 중견국으로서 한국의 국가적 운명을 좌우할 가능성이 크기 때문이다. 다만, 인공지능의 국가전략을 모색하는 과정에서 염두에 두어야 할 점은 인공지능으로 대변되는 미래 기술 패러다임에 대응하는 전략의 강약 조절이다. 인공지능은 글로벌 경제를 강타한 이슈이지만, 그에 대한 대비책은 처한 상황에 따라 다를 수밖에 없으며, 한국의 역량에 맞는 경쟁전략을 고민해야 한다. 사실 이러한 문제제기는 좀 더 넓은 의미에서 본 4차 산업혁명 분야의 경쟁에 임하는 한국의 미래전략에도 적용된다.(김상배, 2017)

참고문헌

- 김대식, 『인간 vs 기계: 인공지능이란 무엇인가』, 동아아시아, 2016.
- 김상배, 『4차 산업혁명과 한국의 미래전략』, 사회평론, 2017.
- 김상배, 『인터넷 권력의 해부』, 한울, 2008.
- 김상배, 『정보혁명과 권력변환: 네트워크 정치학의 시각』, 한울, 2010.
- 김상배 · 황주성, 『소셜 미디어 시대를 읽다: 인터넷 권력의 해부 2.0』, 한울, 2014.
- 레이 커즈와일, 『특이점이 온다』, 김명남 옮김, 김영사, 2007.
- 브루스 슈나이더, 『당신은 데이터의 주인이 아니다』, 이현주 옮김, 반비, 2016.
- 알렉 로스, 『미래산업보고서』, 안기순 옮김, 사회평론, 2016.
- 유발 하라리, 『사피엔스』, 조현욱 옮김, 김영사, 2015.
- 이상욱, 「포스트휴먼시대의 정치사회적 쟁점」, Future Horizon 26, 2015, pp.22-25,
- 이원태, 「인공지능의 규범이슈와 정책적 시사점」, KISDI Premium Report, 2015, pp.15-07.
- 이종관, 「포스트휴먼을 향한 인간의 미래」, Future Horizon 26, 2015, pp.4-9.
- 조성배, 「인공지능 기술의 원리, 현황, 전망」, 기술사회연구회 세미나 발표문, 2015. 10. 16.
- 하원규 · 최남희, 『제4차 산업혁명』, 콘텐츠하다, 2015.
- Ben Goertzel, “Superintelligence: Fears, Promises ad Poten-

tials,” *Journal of Evolution & Technology* 24(2), 2015, pp.55-87.

Bruno Latour, *Reassessing the Social: An Introduction to Actor-network Theory*, Oxford and New York: Oxford University Press, 2005.

David Burke, “Moral Responsibility and Autonomous Machines,” *The End of the Beginning: Life, Society and Economy on the Brink of the Singularity*, Ben and Ted Goertzel, (eds.), Humanity Press, 2015, pp.196-212.

Erika Cudworth and Stephen Hobden, *Posthuman International Relations: Complexity, Ecologism and Global Politics*, London and New York: Zed Books, 2011.

Francis Heylighen and Johan Bollen, “The World-Wide Web as a Super-Brain: from metaphor to model,” *Cybernetics and Systems '96*, R. Trappl (ed.), 1996, pp.917-922.

Francis Heylighen, “Conceptions of a Global Brain: An Historical Review,” *Evolution: Cosmic, Biological, and Social*, 2011, pp.274 – 289.

Francis Heylighen, “Return to Eden? Promises and Perils on the Road to Global Superintelligence,” *The End of the Beginning: Life, Society and Economy on the Brink of the Singularity*, Ben and Ted Goertzel(eds), Humanity Press, 2015, pp.243-306.

Goertzel, *The End of the Beginning: Life, Society and Economy on the Brink of the Singularity*, Ben and Ted Goertzel(eds.), Humanity Press, 2015.

Hugo de Garis, “How will the Artefact War Start?,” *The End of*

the Beginning: Life, Society and Economy on the Brink of the Singularity, Ben and Ted Goertzel(eds.), Humanity Press, 2015, pp.213-241.

Klaus Schwab, *The Fourth Industrial Revolution*, World Economic Forum, 2016.

Michel Foucault, “Governmentality,” *The Foucault Effect: Studies in Governmentality*, Graham Burchell, Colin Gordon and Peter Miller(eds.), The University of Chicago Press, 1991, pp.87-104,

Michel Foucault, *Power/Knowledge: Selected Interviews and Other Writings, 1972-1977*, Pantheon Books, 1980.

Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press, 2014.

인공지능에 대한 이해와 그 문제들

한국사회 인공지능 담론의 비판적 조망

김명호

우리 사회에서 인공지능은 어떻게 이해되고 있는가, 이해의 내용과 틀에 어떤 문제가 있는가, 우리에게 어떤 과제를 제기하고 있는가? 이러한 질문에 답하기 위해, 이 글은 인공지능에 대한 우리 사회의 주요 담론을 비판적으로 분석/조망해 보았다. 세 가지 주요한 갈래로 나뉘는 담론들은 첫째, 인공지능의 산업적 담론(industrial visions), 둘째, 개인적 생존기술의 담론(personal development techniques), 셋째, AI 시대 인간의 우울한 미래를 그리는 비관적 담론(dystopian scenarios)이다. 이들 담론은 첫째, 인공지능 자체에 대한 이야기가 아니라 인공지능이 할 수 있는 또는 할 수 있다고 믿거나 기대하는 역량과 그것이 가지고 올 사회적 과장에 대한 이야기라는 문제, 둘째, 인공지능이 가져올 미래에 주목할 뿐 '지금 여기'에서의 문제를 개선하는 데 인공지능이 무엇

을 할 수 있는 것인지 제대로 담아내지 못하는 문제, 셋째, 낙관과 비관으로 나뉘는 과학과 기술에 대한 전통적 이야기 방식이 소재만 달리한 채 반복되는 문제, 넷째, 유행과 같은 양상을 띠면서 담론 내용과 수준의 천박함을 드러내고 있는 문제, 다섯째, 사실상 시민은 배제된 채 전문가 위주의 폐쇄적 양상으로 전개되고 있는 문제 등을 품고 있다.

이 같은 정황은 인공지능 분야와 연관되어 있는 다양한 주제들에게 많은 과제를 던져준다. 첫째, 컴퓨터나 로봇공학 전문가들은 인공지능에 대한 비판적 이해를 증진시키는 대중적 노력을 다양한 방식으로 전개해야 한다. 또 지금까지 우리 사회가 수행해온 인공지능 연구의 내용과 경과, 주제, 연구성과와 수준 등을 정확하게 기록/진단/평가하고 그것을 사회적으로 공유토록 해야 한다. 둘째, 과학기술 분야 시민사회 단체 역시 인공지능 문제에 대한 사회적 인식을 높일 수 있는 다양한 프로그램을 진행, 다가오는 변화의 시대에 능동적으로 대처하는 시민 집단을 키워야 한다. 셋째, 사회과학 분야 연구자들은 인공지능과 연관되어 있는 현재와 미래의 무수한 사회적 문제에 대해 냉정한 시각에서 진단과 분석, 대안을 제시하는 노력을 기울여야 한다. 인공지능과 관련해 난무하는 허황한 언술들을 제어하고 적절한 사회적 대응을 모색하는 데 이는 필수적 과제이다. 넷째, 정책 당국자들은 인공지능 분야의 지도적 주체로 나서는 것이 아니라, 관련 연구집단을 지원하는 후견인, 그리고 다양한 연구집단이나 조직/기관들 간의 협력 시스템 구축을 지원하거나 매개하는 조정자 역할을 수행해야 한다. 다섯째, 인공지능과 연관되어 있는 기업들은 상업적 성공, 신수종 비즈니스 같은 응용적 차원으로 인공지능에 접근하는 과장된 기대를 버려야 한다. 단기적 집착은 기업이 하청업자 수준에 머무르는 지름길이며 심한 경우 장기적 패착을 낳기 때문이다. 보다 본질적인 문제를 탐색하는 작업에 나서거