AI기본법안과 '제3의 길': 진흥과 규제 사이*

강지원 김·장 법률사무소 외국변호사(미국 뉴욕주)

I. 서론

우리나라에서 인공지능(AI)이 처음 폭발적인 대중적 관심의 대상이 되었던 것은 2016년 3월 알파고와 이세돌의 딥마인드 챌린지 매치일 것이다. 무한에 가까운 경우의 수를 계산해야 하는 바둑에 있어서 아직 인공지능이 인간을 넘어서기엔 시기상조일 것으로 생각되었으나, 결과는 예상 밖으로 알파고가 압도적인 승리를

* 본고의 작성과정에서 많은 도움을 주신 김·장 법률사무소의 한기웅 변호사님께 감사드립니다.

거두며 인공지능의 놀라운 발전 속도에 많은 관심이 모아졌다. 그러나 바둑과 같이 제한된 규칙 내에서 작동하는 분야 외에 인공지능이 본격적으로 일상의 많은 분야에 활용될 수 있을지에 대해서는 여전히 불확실해 보였다.

그런데 2022년 말 챗GPT(ChatGPT)가 세상에 공개되면서, 이제는 인공지능이 일상을 바꿔놓을 정도의 파급력을 갖게 될 것인지 여부 자체보다는 그 속도와 방향에 관심이 집중되고 있다. 소위 대규모 언어 모델(Large Language Model, LLM)이 급속도로 발전하면서 언어를 활용한 인간의 많은 일들이 인공지능에 대체될 수 있게 되었고, 사회 각 분야에서 인공지능이 불러올 변화에 대한 장밋빛 전망이 쏟아졌다. 2024년 하반기 국내에서는 딥페이크 영상을 활용한 성범죄에 대한 각종 피해사례가 밝혀지며, 인공지능이 앞으로 우리사회에 제기할 위험에 대한 경각심도 한층 높아졌다.

이처럼 최근 인공지능이 인간 사회에 미치는 영향의 양면적 효과를 피부로 체감하게 되면서, 한편으로는 인공지능산업의 진흥과 발전을 지원하고, 또 다른 한편으로는 인공지능 이용자들을 보호하기 위하여 인공지능 시대의 기본 법제로서 인공지능(AI)기본법을 제정해야 한다는 목소리가 높아졌다. 이미 유럽은 2020년 유럽연합(EU)집행위원회가 〈인공지능 백서(White Paper on Artificial Intelligence)〉를 발간한 이래로 다양한 논의를 거쳐 2024년 8월 2일 「EU AI법(EU AI Act)」이 최종 발효되었다. 해당 법은 허용되지 않는 위험 AI(Unacceptable Risk AI), 고위험 AI(High-risk AI), 제한된 위험 AI(Limited risk AI), 최소위험 AI(Minimal risk AI) 등을 구분하여 차등적으로 규제하는 소위 위험기반 접근 방식(risk-based approach)을 취하고 있는데, 현재 세계 유일의 포괄적인 인공지능 규율 법제로서 세계 각국의 관련 입법 과정에서 가장 많이 참고하는 선례가 되고 있다.

다만 이런 위험기반 접근 방식에 관한 비판 역시도 제기되고 있는 것이 사실이다. 특히 미국을 중심으로 EU의 규제체계가 지나치게 경직적이며, 산업별로 맥락에 맞는 규제를 유연하게 적용하는 방식이 필요하다는 의견이 제기되고 있다. 혹자는 미국이 소위 빅테크 기업들을 앞세워 인공지능 산업에서 우위를 점하고 있는 상황으로 인해 미국과 유럽 모두 각자의 이해관계에 맞는 방식을 내세우고 있다고 평가한다.

우리나라도 AI기본법 마련에 속도를 높이고 있어 조속한 시기에 제정안 통과가

점쳐지고 있다. 다만 산업계에서는 지나친 규제로 인한 인공지능 산업 발전의 저해를 우려하고, 시민사회에서는 이용자 보호에 미흡한 점들을 지적하며 각기 다른 관점에서 균형 있는 법안 마련을 위한 다양한 의견들을 개진하고 있다. 이런 상황 속에서, 제21-22대 국회의 AI기본법 제정 논의를 돌아보고, 현재 국회 소관 상임위원회를 통과한 인공지능기본법 제정안의 내용을 살펴보고자 한다.

Ⅱ. 기존의 AI기본법 입법 논의 동향 및 경과(제21-22대 국회를 중심으로)

1) 입법 경과

지난 제21대 국회에서는 인공지능 관련 법률안이 총 9개 발의되었다. 관련 법률을 소관하는 과학기술정보방송통신위원회(이하 '과방위')에서는 2022년까지 발의된 7개 법률안을 병합하여 심사한 후, 이를 정리하여 과방위 위원장 명의로 대안을 제시하기로 하였으나, 비공개로 수정·보완 중인 상태에서 21대 국회의 임기만료로 최종 폐기되었다.

이후 제22대 국회가 개원한지 약 반년 밖에 경과하지 않은 현재(2024년 12월 9일 기준) 발의된 AI기본법 관련 법안은 총 20개로, 여·야 모두 AI기본법 마련의 필요성에 공감하는 모양새다. 국회는 신속히 AI기본법을 제정하는 것을 목표로 속도를 내고 있는데, 2024년 9월 24일 AI기본법 제정을 위한 국회 공청회를 개최하여 시민사회 등의 목소리를 청취하였으며, 여·야당 대표 간 회동에서 AI기본법의 조속한 추진에 합의하였다고 알려졌다. 이에 따라 발의된 19개의 법안(1개 법안은 법안심사소위 통과 후 발의)을 병합한 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법안」(이하 'AI기본법안')이 국회 과방위 제2소위(정보통신방송법안심사소위원회) 심사를 통과하였으며, 2024년 11월 26일 과방위 전체회의에서 의결되어 현재 법제사법위원회 심사와 본회의 의결만을 앞두고 있다.

2) 논의 과정에서 주목된 법안

위 20개 의원안 모두 각자의 특색이 있으나, 논의에서 주목해 볼 수 있는 법안으로는 정점식 의원안, 권칠승 의원안, 최민희 의원안을 꼽을 수 있다. 정점식 의원안은 22대 국회 개원 후 2번째로 발의된 AI기본법안으로, 여당인 국민의힘이 당론 발의하였다는 점에서 가장 중요한 법안으로 여겨진다. 앞서 언급한 2024년

11월 21일 국회 과방위 제2소위 심사 당시 과학기술정보통신부(이하 '과기부')가 기존에 발의된 법안들을 통합하여 마련한 병합안을 중심으로 심사가 이루어졌는데, 해당 병합안은 정점식 의원안을 기초로 공청회 논의 내용 및 타 법안 내용 일부를 수용하여 만들어졌다. 후술하는 바와 같이 정점식 의원안의 기본적인 체계와 핵심 내용은 국회 과방위를 통과한 AI기본법안에 상당 부분 반영되었는데, 한 가지 중요한 차이점은 제재 규정의 유무라고 할 수 있다. 정점식 의원안은 고위험 인공지능 등에 대해 여러 준수의무를 부과하면서도 그 의무 위반에 대하여 행정제재나 형사처벌 규정이 없어 실효성에 의문을 제기하는 견해가 제기되었는데, 병합안이 법안심사소위를 통과하는 과정에서 제재 규정이 추가되었다.

다음으로 권칠승 의원안(2024년 7월 4일 발의)의 경우 제22대 국회 개원 초창기에 발의된 법안 중 유일하게 금지된 인공지능을 규정하고, 사업자에 대한 벌칙 규정을 마련하는 등 가장 규제적인 성격을 띤 법안이라는 점에서 의미가 있다. 해당 법안은 "인간의 존엄과 가치, 인류의 평화와 안전에 대한 심각한 침해나 위협이 명백하다고 인정되어 개발과 이용이 금지된 인공지능"을 금지된 인공지능으로 규정하면서도 그 세부 기준 및 유형은 대통령령으로 정하게 하였으며, 아주 예외적인 경우를 제외하고는 해당 인공지능의 개발·이용을 전면 금지하였다. 이는 금지된 인공지능을 규정한 「EU AI법」의 영향을 받은 것으로 생각되는데, 현재 과방위를 통과한 AI기본법안에는 인공지능산업 발전의 잠재력을 저해하지 않아야 한다는 우려를 감안하여 금지된 인공지능 규정이 반영되지 않았다. 권칠승 의원안은 또한 인공지능사업자의 다양한 의무 위반에 대하여 형사처벌 규정을 두었는데, 강력한 형사처벌 규정으로 인해 인공지능산업의 발전이 위축될 것이라는 우려의 견해와 인공지능의 안전성·신뢰성 보장을 위해 법적 구속력이 필요하다는 의견이 존재했다.

마지막으로 최민희 의원안의 경우 상대적으로 최근에 발의되었는데(2024년 11월 14일), AI기본법안의 소관 상임위원회인 과방위원장이라는 점에서 많은 관심을 받았다. 최민희 의원안은 고위험 인공지능에 관하여 각종 규제를 부과하고 있는데, 제공자와 운영자를 구분하여 차별화된 의무를 부과하고 있는 점이 눈에 띈다. 후술하는 것처럼 과방위를 통과한 AI기본법안도 인공지능개발사업자와 이용사업자를 구분하여 정의하고 있는데, 법안에서는 개발·이용사업자 별로 의무를 구분하고 있지는 않으나, 향후 시행령 등을 통해 사업자의 유형에 따라 의무가 구분될 가능성이 있다. 최민희 의원안의 내용이 참고가 될 여지가 있다는 점에서 주목할 만하다.

Ⅲ. 국회 과방위를 통과한 인공지능기본법안의 주요 내용

1) 적용대상 및 산업 진흥 규정

AI기본법안의 적용대상 범위 설정은 '인공지능'의 개념을 어디까지로 볼 것인지, 그리고 인공지능 관련한 사업자의 범위를 어떻게 설정할 것인지와 관련된다. 적용대상의 범위와 관련하여 주목할 부분은 본 법안의 성격이 진흥과 규제를 동시에 아우른다는 점이다. AI기본법안은 하나의 법안에서 진흥 및 지원규정의 적용대상은 가급적 넓게 확대하는 한편 과도한 규제로 인한 산업발전 저해 우려는 최소화하기 위해, '인공지능'과 '인공지능시스템'을 각각 별도로 정의하여 적용대상을 달리 하고 있다.

'인공지능'은 "학습, 추론, 지각, 판단, 언어의 이해 등 인간이 가진 지적 능력을 전자적 방법으로 구현한 것"으로 다소 광범위하게 정의하여, 정부 지원의 대상인 '인공지능기술' 및 '인공지능산업'과 관련된 사업자의 범위 역시 폭넓게 규정된다. 이에 반해, 규제의 대상이 되는 '인공지능시스템'은 「EU AI법」과 유사하게 "다양한 수준의 자율성과 적응성을 가지고 (…) 예측, 추천, 결정 등의 결과물을 추론"하는 것으로 정의하고 있는데, 이는 주로 머신러닝 기술 등을 채택한 인공지능을 대상으로 하여 '인공지능'에서 그 범위를 좁힌 것으로 볼 수 있다. 인공지능산업의 후발주자인 우리나라의 현실을 고려하여, 이용자 보호를 위해 반드시 필요한 규제

는 적용대상을 좁혀 법안에 담되 인공지능기술 개발과 확산을 위한 정부 지원의 적용대상은 넓어질 수 있도록 한 입법자의 취지로 여겨진다. 정부는 인공지능기술의 표준 마련, 인공지능산업의 "쌀"로 불리는 학습데이터의 수집·활용 지원, 인공지능 집적단지 지정, 데이터센터 구축의 활성화 등 인공지능산업 발전에 필요한 각 방면의 지원 근거를 본 법안을 통해 확보하게 되었다. 본 법안의 적용대상인 수범자의 경우, 기존의 정점식 의원안은 "인공지능사업자"라는 단일한 주체를 "인공지능과 관련된 사업을 하는 자"로 매우 포괄적으로 규정하여, 인공지능의 개발·유통·활용 등에 따른 수범자 별 다양한 역할과 책임을 구분하지 않고 있다는 비판이 제기된 바 있다. AI기본법안은 이를 보완하여 "인공지능사업자"의 하위 개념으로 "인공지능개발사업자"와 "인공지능이용사업자"를 구분하여 정의하고 있는 점이 특징이다. 다만, 인공지능개발사업자는 인공지능 모델 및 학습데이터 등에 관한 이해도와 통제 권한을 보유하고 있는 반면, 인공지능이용사업자는 다양한 산업군에서 그 이용 목적에 맞게 인공지능을 활용할 뿐 아니라 최종이용자와 접점에서 제품·서비스를 제공한다는 점에서 각자의 역할과 성격에 부합하는 의무 구분이 필요할 것이다. 법률에는 이러한 세부적인 구분이 크게 드러나지 않고 있는데, 향후 시행령과 고시를 통해 구체화될 내용으로 보인다.

2) 고영향 인공지능의 범위 및 준수의무

한편, AI기본법안의 규제 내용 중 가장 적용범위가 넓고 향후 업계에 중요한 영향을 미칠 것으로 보이는 것은 고영향 인공지능 규제라고 할 수 있다. 법안은 "생명, 신체의 안전, 기본권에 중대한 영향을 미치거나 위험을 초래할 우려가 있는 인공지능 시스템"을 고영향 인공지능으로 정의하여, 기본권에 미치는 영향을 사전에 평가하고 그 개발·이용에 있어 안전성과 신뢰성을 확보할 수 있는 조치를 취하도록 하는 등, 이용자에게 미칠 위험을 사전·사후적으로 관리하도록 의무화하고 있다.

AI기본법안의 기초적인 틀을 제공한 정점식 의원안에서는 '고위험' 인공지능이라는 용어를 사용했으나 법안심사소위에서 병합안이 마련되는 과정에서 '고영향' 인공지능으로 변경되었는데, 이는 '위험'이라는 단어가 주는 부정적 의미를 피하고 산업에 미치는 중요성을 부각하는 보다 중립적인 용어를 채택하는 것이 바람직하다는 이해민 의원안의 입장이 반영된 것으로 알려진다. 자신이 개발·이용하는 인공지능시스템이 고영향에 해당되는지 여부는 과기부장관이 제정하는 고시로 마

련되는 일정 기준을 참고하여 인공지능사업자가 자율적으로 판단하도록 하되, 고영향 해당 여부에 대한 확인을 과기부에 요청할 수 있도록 하고 있어 실무상 규제 불확실성 해소를 위해 과기부에 확인을 요청하는 사례가 많을 것으로 보인다.

고영향 인공지능의 유형은 에너지, 의료기기, 교통수단 등 생명·안전과 관련된 유형, 범죄 수사·체포, 채용, 학생평가 등 기본권과 연관된 유형 등으로, 법률에 열거된 유형 뿐 아니라 시행령을 통해 추가로 규정할 수 있도록 하고 있어 향후 고영향 인공지능의 범위가 확대될 것으로 예상된다. 고영향 인공지능을 이용한 제품 또는 서비스를 제공할 경우 이용자에게 반드시 그러한 사실을 사전에 알려야 할 뿐 아니라, 기본권에 미치는 영향을 사전에 평가하기 위해 노력해야 한다. 이 때 고영향 인공지능에 기반하여 운용된다는 사실에 대한 사전고지는 그 이행이 필수적인데 반해, 기본권에 미치는 영향평가 수행은 「EU AI법」과 달리 '노력'할 의무만 부과하고 있다는 점에서 차이점을 보인다. AI기본법안의 성격이 규제는 필요 최소한의 내용만 담고 산업의 진흥 및 기술 발전에 보다 역점을 두고 있음을 엿볼 수 있는 부분이다.

고영향 인공지능의 개발·이용사업자(인공지능사업자)는 그 안전성·신뢰성을 확보하기 위한 조치를 이행할 의무를 부담하는데, 「EU AI법」과 마찬가지로 본 법안의 핵심적인 규제사항이 될 것으로 보인다. 인공지능사업자는 위험관리방안의 수립, 사람의 관리·감독, 이용자 보호 방안의 수립, 안전성·신뢰성 확보조치 내용을 확인할 수 있는 문서의 작성·보관 등의 의무를 이행해야 한다. 이러한 조치는 학습데이터의 편향성에 따라 발생하는 인공지능 결과물의 특정 집단에 대한 사회적 차별 방지, 인공지능 결과물 도출의 불투명성에 따른 책임 소재의 파악 등을 개선하기 위해 요구되는 최소한의 관리조치라고 볼 수 있다.

안전성·신뢰성 확보조치 중 특히 주목할 내용은 인공지능이 도출하는 결과물에 대한 설명의무이다. 머신러닝(machine learning)과 같은 인공지능기술은 사람이 기존에 입력한 규칙에 따라 결과물을 산출하지 않고 스스로 목표 달성과 결과물 도출을 위한 최적의 규칙을 찾아내는 특성이 있기에, 그 결과물에 대한 설명가능성(explainability)을 기술적으로 가능한 범위 내에서 보장하는 것이 중요하다. 인공지능기본법안은 인공지능이 도출한 결과 뿐 아니라 도출과정에서 사용된 주요 기준 및 학습데이터에 대한 설명 방안 마련을 의무화하고, 그러한 설명을 요구할 수 있는 권리(설명요구권)까지 부여하고 있다.

3) 생성형 인공지능 등에 대한 별도의 규제

사람의 지시에 따라 글, 소리, 그림, 영상 등을 만들어내는 생성형 인공지능은 산업계와 문화·예술계 등 전방위적으로 그 응용이 확산되고 있기에, 「EU AI법」뿐 아니라 AI기본법안에서도 일정한 법적 의무를 담고 있다. 다만, 생성형 인공지능은 그 범용적 특성상 그 자체만으로 위험도가 높은 고영향으로 분류하기 어려운 속성이 있기에, 「EU AI법」에서도 고위험으로 분류하지는 않고 있으며, 기본법안 역시 유사한 접근을 취한다. 생성형 인공지능의 개발·이용사업자는 고영향 인공지능과 마찬가지로 제품 또는 서비스가 생성형 인공지능 기반으로 운용된다는 사실을 이용자에게 알려야 할 뿐 아니라, 그 결과물이 인공지능에 의해 생성되었다는 사실을 표시할 의무를 추가로 부담한다. 특히 최근 딥페이크로 인한 여러 사회적 폐해가 현실화됨에 따라, 실제와 구분하기 어려운 딥페이크 영상 등을 인공지능으로 만들어낼 경우 실제 이미지가 아닌 가상의 생성물임을 고지·표시하도록 하고 있다.

마지막으로, 학습에 사용된 누적 연산량이 일정 기준을 넘는 인공지능의 경우 별도의 영향평가나 고영향 여부 확인을 거치지 않고 위험을 식별·평가·완화하고 안전사고를 모니터링·대응하기 위한 위험관리체계를 구축하도록 의무화하고 있는 점이 특징이다. 「EU AI법」에서는 이를 고영향 범용 AI(high-impact general purpose AI)로 구분하여 유사한 규제를 적용하고 있는데, AI기본법안은 '고위험' 대신 '고영향' 개념을 이미 적용하고 있어 구분을 위해 용어를 달리 하고 있는 것으로 보이는데, 규제 적용기준을 인공지능 성능의 주요 척도인 학습에 사용된 연산량으로 설정하고 위험관리 의무를 두고 있는 점은 대체로 유사하다. 적용기준인 학습에 사용된 연산량은 인공지능기술의 빠른 발전속도를 고려해 시장 변화에 탄력적으로 대처할 수 있도록 법률로 정하지 않고 시행령에 위임하고 있다. 이러한 고성능 인공지능은 각 분야에 활용됨에 따른 경제사회적인 파급효과가 상당할 수 있다는 점에서 위험관리 의무의 이행결과를 과기부에 별도로 제출하도록 하고 있다.

Ⅳ. 마치며

지난 제21대 국회부터 장기간 이어져 온 인공지능법 제정의 화두는, 인공지능산업과 기술의 발전을 위한 육성을 적극 지원하는 동시에, 안전하고 신뢰성 있는 인공지능의 이용을 보장하는 균형 있는 입법의 추구였다. 산업진흥에 좀 더 방점이

놓여 있던 것으로 보였던 초기의 법안은 입법 논의 과정에서 「EU AI법」의 제정과 딥페이크 등 인공지능 위험의 현실화 등을 반영하여 규제 내용이 점차 강화되는 방향으로 정리된 것으로 보인다.

다만 규제의 강도와 적정성을 판단할 수 있는 실질적인 내용은 대부분 시행령과 고시 등 하위 법령에 위임되어 있어, 현재는 다소 추상적인 수준의 준수의무만 법률로 규정하고 있다고 할 수 있을 것이다. 현재 10개 유형만 명시된 고영향 인공지능의 범위가 시행령을 통해 얼마나 확대될 것인지, 위험관리의무 이행을 위해 필요한 기술적·관리적 조치의 내용과 범위가 어느 정도 수준일지, 그리고 합리적 수준의 설명가능성 보장은 어디까지 인정될 수 있을 것인지 등은 현 시점에서 쉽게 예측하기 어렵다.

한편, 안전성·신뢰성 보장에 관한 모든 사항을 정부가 규율하지 않고 일정 부분 민간의 자율규제 영역에 맡겨 놓은 점도 본 법안이 우리나라 현실을 적절히 반영하여 「EU AI법」과 차별화를 도모한 부분으로 보인다. 인공지능윤리위원회 등 조직 내 인공지능 개발·이용의 컨트롤 타워는 각 기업·기관의 상황에 맞게 자율적으로 운영 여부를 결정하면서도, 운영할 경우 그 구성·기능에 있어 최소한의 공통분모를 둘 수 있도록 한 점이 그 예이다. 안전성·신뢰성 검·인증 역시 「EU AI법」은 시장 출시 전 필수적인 법적 의무(적합성 평가)로 규정하고 있지만, AI기본법안에서는 민간 자율의 영역으로 두고 정부는 민간의 검·인증 활동이 원활히 작동할 수 있는 기반 조성에 주력하도록 하고 있는 점도 차이를 보인다.

결론적으로, AI기본법안은 EU, 미국과 차별화된 "제3의 길", 진흥과 규제 간의 균형이라는 우리나라 인공지능산업의 현실을 반영한 내용으로 구성되어 있다는 점에서 긍정적인 평가를 내릴 수 있을 것으로 보인다. 미국 빅테크 기업의 시장 잠식을 견제하기 위한 강도 높은 규제법 제정을 선택한 EU, 자국 빅테크 기업의 육성을 우선하면서 국가안보 차원의 규제 중심으로 범위를 설정하는 미국 등 각국이 자국의 인공지능산업을 명시적으로 우선하는 입법과 정책에 속도전을 내는 상황에서, 더 늦기 전에 AI기본법안의 필요성과 방향에 대한 사회적 합의가 도출된 것은 고무적인 일이다. 국회에서의 남은 입법과정, 그리고 규제의 핵심적인 내용이 담길 하위법령 마련 과정에서도 산업계와 이용자, 그리고 전문가 집단 등의 다양한 의견이 적절히 조화된 결과물이 나올 수 있기를 기대한다.