pwc

AI 기본법 시행, 기업은 무엇을 준비해야 하는가?

September 2025

Table of contents

들어가며

AI를 활용한 기업의 미래 성장 동력 창출이 기업 생존의 필수요소임은 그 누구도 부정할 수 없게 되었으며, 이러한 기술 중심의 변화 속에서, AI의 책임 있는 활용을 보장하기 위한 법적·제도적 장치도 빠르게 정비되고 있습니다.

유럽연합의 EU AI Act에 이어, 대한민국은 세계에서 두번째로 포괄적 AI 규제 체계를 담은 'AI 기본법' 을 제정하였으며, 시행까지 약 3개월이 남은 가운데, 법의 구체적인 집행을 위한 시행령과 5개 분야별 가이드라인의 배포 일정이 확정 되었습니다. 정부는 오는 10월 입법예고를 거쳐 연내 제정을 마무리할 방침입니다.

이에 따라 기업은 단순한 규제 대응을 넘어, 자사 AI 서비스의 특성과 활용 목적에 맞는 실용적이고 효율적인 관리체계를 선제적으로 수립해야 할 상황에 놓여 있습니다.

본 보고서는 AI 기본법의 규제적 측면과 기업이 준비해야 할 대응 전략, 그리고 실무 중심의 AI 관리체계 구축 방안을 중심으로, 법 시행 이전에 기업이 고려해야 할 핵심 사항들을 정리합니다. AI 기술을 활용한 혁신을 지속하면서도 법적 책임을 다할 수 있는 전략적 방향성을 모색하는 데 도움이 되기를 바랍니다.

Ⅰ. AI 기본법 주요 동향

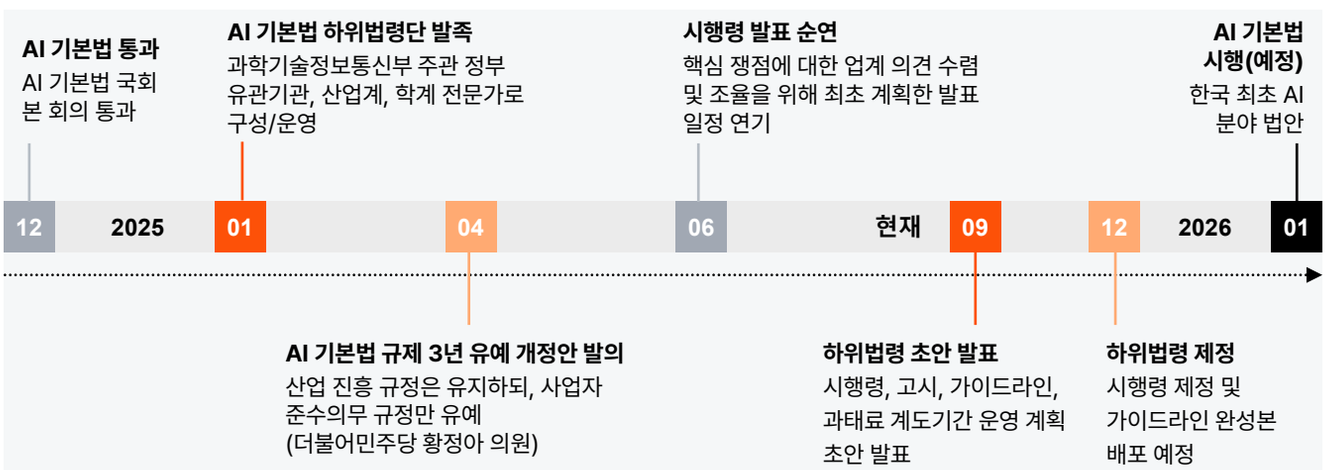
01. AI 기본법이란 무엇인가?

AI 기본법의 제정 경과

「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」 (이하 'AI 기본법')이 2026년 1월 22일부터 시행됩니다. AI 기본법은 AI 기술의 건전한 발전과 사회적 신뢰 기반 조성을 목표로 합니다.

2024년 12월 국회 본회의에서 가결되었으며 시행에 앞서 하위법령 및 가이드라인을 제정 중입니다.

입법 및 규제 동향



출처: PwC Analysis

AI 관련 법안 제정은 전 세계에서 우리나라가 EU에 이어 두 번째입니다. EU의 「AI Act」가 2026년 8월에 전면 발효될 것임을 감안하면, 실제 AI에 대한 전면적인 법 적용은 AI 기본법이 세계 최초입니다. 이러한 AI 기본법은 국내 AI 산업의 진흥과 함께 안전·신뢰 기반 구축이라는 규제적 요소도 포함하고 있습니다.

AI 기본법 주요 규정 사항

국가 AI 정책 거버넌스 확립	AI 산업 육성 지원	안전신뢰 기반 조성
▪ 국가AI위원회를 설치하고, AI 정책을 심의· 의결하도록 법률상 근거 마련	▪ AI 개발·도입·활용 전영역에서 정부의 지원 사항 마련	▪ AI의 활용을 촉진하고, 잠재적 위험으로부터 국민을 보호
▪ 국가AI위원회 설치·기능, 분과· 특별위원회, 지원단 설치 등	▪ AI R&D 지원, 기술 표준화, 학습용 데이터 구축, AI기술 도입·활용 지원, 법제도 개선 등	▪ AI 윤리원칙, AI 윤리위원회 설치 ▪ 투명성·안전성 확보 의무, 고영향 AI 확인·사업자 책무 등

출처: AI 기본법, PwC Analysis

AI 기본법의 주요 키워드 '고영향 AI'

AI 기본법은 학습, 추론, 지각, 판단, 언어의 이해 등 인간이 가진 지적 능력을 전자적 방법으로 구현한 것을 AI로 규정합니다.

이때, AI 기본법의 가장 큰 특징은 '고영향 AI'를 규정한 것입니다. 고영향 AI는 에너지 공급, 먹는 물 생산 등 10개 영역에서 사용되는 AI 중에서 '사람의 생명, 신체의 안전 및 기본권에 중대한 영향을 미치거나 위험을 초래할 우려'가 있는 AI를 의미합니다.

고영향 AI 사업자는 고영향 AI 사전 검토, 사전 고지, 안전성·신뢰성 확보조치, 안전·신뢰 검·인증, 영향평가 등의 의무를 이행해야 합니다.

AI 기본법상 고영향 AI와 고영향 AI 사업자 의무

고영향 AI 명시

	근거조항(법)	유형
1	「에너지법」	에너지의 공급
2	「먹는물관리법」	먹는 물의 생산 공정
3	「보건의료기본법」	보건의료 제공·이용체계의 구축 및 운영
4	「의료기기법」	디지털 의료기기 개발 및 이용
5	「방사능방재대책법」	핵물질 원자력 시설의 안전한 관리 운영
6	범죄 수사, 체포 업무 위한 생체인식정보 분석, 활용	
7	채용, 대출심사 등 개인 권리·의무관계에 직결된 평가, 판단	
8	「교통안전법」	교통수단, 교통시설, 교통체계의 주요 작동, 운영
9	공공서비스 제공 및 국민에 대한 국가기관 의사결정	
10	「교육기본법」	유아·초등 교육 및 중등 교육에서의 학생 평가

고영향 AI 사업자 의무

사전 검토 의무	인공지능사업자는 고영향 인공지능에 해당하는지 여부를 사전에 검토해야 하며, 필요한 경우 과학기술정보통신부 장관에게 확인 요청이 가능함 (제33조)
사전 고지 의무	고영향 인공지능을 이용한 제품·서비스를 제공하는 경우 인공지능이용사업자에게 사전고지의무가 부과됨 (제31조 제1항)
안전성·신뢰성 확보 조치 이행 의무 의무	고영향 인공지능 또는 이를 이용한 제품·서비스를 제공하는 경우 일정한 안전성·신뢰성 확보조치 이행의무가 부과됨 (제34조)
고영향 AI 안전·신뢰 검·인증 권고	인공지능사업자가 고영향 AI를 제공하는 경우 사전에 안전성·신뢰성에 대한 자율적인 검·인증을 받도록 노력해야 함 (제30조 3항)
기본권에 미치는 영향 평가 권고	인공지능사업자는 고영향 AI를 이용할 경우 사람의 기본권에 미치는 영향을 평가하기 위해 노력해야 함 (제35조 1항)

출처: AI 기본법, PwC Analysis

AI 기본법의 주요 키워드 '투명성, 안전성'

AI 기본법은 AI 제품·서비스를 개발·제조·제공하는 '인공지능산업'과 관련된 사업을 하는 자를 '인공지능사업자(이하 AI 사업자)' 로 정의합니다. AI 사업자는 투명성 의무, 고성능 AI의 안전성 확보 의무를 이행해야 합니다.

AI 사업자의 의무

투명성 의무	AI 사업자는 자신의 제품·서비스가 고영향 또는 생성형 AI에 기반하여 운용될 경우, 그 사실을 이용자에게 사전에 고지해야 함 (제31조1항)
	AI 사업자는 자신이 제공하는 제품·서비스의 결과물이 생성형 AI에 의해 생성된 경우, 그 사실을 표시해야 함 (제31조2항)
	AI 사업자는 AI를 이용하여 실제와 구분하기 어려운 가상의 결과물(예: 딥페이크)을 제공하는 경우, 해당 결과물이 AI에 의해 생성되었다는 사실을 이용자가 명확하게 인식할 수 있는 방식으로 고지 또는 표시해야 함 (제31조3항)
고성능 AI의 안전성 확보 의무	AI 사업자 중에서 학습에 사용된 누적 연산량이 큰 대규모 AI를 사용하는 경우, 그 안전성 확보를 위하여 △위험의 식별·평가 및 완화 조치, △ 안전사고 모니터링 및 위험관리체계 구축, △이행 결과의 과학기술정보통신부장관 제출 등을 이행해야 함 (제32조)

출처: AI 기본법, PwC Analysis

이러한 의무를 이행해야 하는 AI 사업자는 고영향 AI, 생성형 AI, 학습에 사용된 누적 연산량이 큰 고성능 AI를 사용하는 사업자입니다.

투명성 확보 관련 사전고지 미이행, 일정 기준 이상 해외사업자의 국내대리인 미지정, AI 기본법 위반에 따른 시정명령 미이행 시 3천만원 이하의 과태료가 부과됩니다.

02.
AI 기본법의 쟁점 및 하위 법령 정비 동향

AI 기본법의 쟁점

AI 기본법을 적용하거나 해석해야 하는 당사자는 어떤 AI가 사람의 생명, 신체 안전, 기본권에 중대한 영향을 미칠 가능성이 있는지를 스스로 판단해야 합니다. 그러나 현재의 법 조항만으로는 이를 일관되게 해석하기 어렵습니다.

또한 제2조 4항은 고영향 AI의 적용 분야 중 10개는 구체적으로 명시하면서도, 마지막 하나는 '그 밖에 사람의 생명·신체의 안전 및 기본권 보호에 중대한 영향을 미치는 영역으로서 대통령령으로 정하는 영역'이라고 규정하고 있습니다. 대통령령이 제정되기 전까지는 이 영역을 예측하기 어렵습니다.

고영향 AI 사업자 의무의 구체적 이행 방안과 투명성, 안전성 확보 의무에 대한 구체적인 이행 방식 역시 AI 기본법에서는 구체적인 내용을 명시하지 않고 있으며 시행령과 고시, 가이드라인에 위임되어 있습니다.

하위 법령 마련 추진 경과

과학기술정보통신부는 AI 기본법 시행을 앞두고 2025년 1월부터 하위법령을 정비하고 있습니다. 2025년 9월 하위법령 초안을 마련하였고 관계 부처 및 이해관계자 설명 및 의견 수렴을 거쳐 2025년 12월까지 시행령 및 고시를 마련하고 가이드라인을 배포할 예정입니다.

AI 기본법 하위법령 정비 주요 사항 및 추진 일정

AI 기본법 하위 법령 정비 주요 사항	▪ 인공지능 기본법 시행령 및 시행규칙 ▪ 고영향 인공지능 기준과 예시에 관한 방침(가이드라인)(제33조) ▪ 고영향 인공지능 사업자 책무 방침(가이드라인)(제34조) ▪ 인공지능 안전성 확보 의무 고시(제32조) ▪ 인공지능 영향평가 방침(가이드라인)(제35조) ▪ 인공지능 투명성 확보 의무 방침(가이드라인)(제31조)
추진 일정	▪ 9월 2~4주: 관계부처, 이해관계자 대상 설명 및 의견 수렴
	▪ 10~11월: 입법 예고, 규제 심사, 법제처 심사 등 행정 입법 절차 진행
	▪ 12월: 시행령, 고시 마련, 가이드라인 공개

출처: AI 기본법, PwC Analysis

하위 법령 주요 내용

하위 법령 초안은 AI 기본법의 적용 범위와 AI 육성을 위한 지원 대상 및 기준, 투명성·안전성 확보 방법, AI 영향 평가 등 법에서 위임하는 기준·절차에 대한 구체적인 사항을 다루고 있습니다. 특히 안전·신뢰 기반 조성을 위한 사항은 고시, 가이드라인을 통해 구체화될 예정입니다.

AI 기본법 하위 법령 초안 주요 내용

적용 범위	▪ 국방·국가 안보 목적으로만 이용되는 AI는 AI 기본법 적용 대상에서 제외
거버넌스	▪ 국가AI위원회의 명칭을 '국가AI전략위원회' 로 변경 ▪ 국가AI전략위원회의 기능에 부처간 정책 조정, 이행점검, 성과관리 사항 추가 ▪ 최고AI책임자(CAIO) 및 CAIO협의회 운영 근거 마련
AI 산업 육성	▪ R&D, 데이터 구축, AI도입·활용 등 산업 육성을 위한 정부 지원 근거 규정을 마련

출처: PwC Analysis

AI 기본법 하위 법령 초안 주요 내용

AI 사업자 의무	▪ **(투명성)** 생성형·고영향 AI 이용자에 대한 사전고지 및 결과물 표시(워터마크) 의무를 부여하며, 의무 이행 방법과 예외*를 시행령에 규정 *▲약관·UI 등을 활용한 사전고지 인정, ▲비가시적 워터마크 인정, ▲딥페이크 결과물에 대해선 이용자의 연령·신체적 조건 등을 고려하여 고지·표시, ▲사업자 내부 업무용이거나, 생성형·고영향 AI 기반이 명백한 경우 투명성 의무 면제 등
	▪ **(안전성)** 누적 학습량이 일정 수준 이상인 고성능AI는 위험 완화 등 안전 확보 의무를 부담하며, 대상 AI의 기준* 및 의무 이행 방식** 등을 시행령·고시에 반영 * 누적 학습량이 10의 26제곱 부동소수점 연산 이상인 AI시스템 중 고도화된 기술을 적용한 AI시스템이며, 사람의 생명, 신체의 안전, 기본권에 광범위하고 중대한 영향을 미칠 수 있는 시스템 ** ▲위험식별(방법론 마련, 기능오류·데이터편향 등 위험요소 확인), ▲평가(지표 마련, 독립된 위험평가팀 활용), ▲완화(우선순위·실행가능성을 고려한 완화조치, 긴급대응계획 수립) 등
	▪ **(고영향 AI)** 고영향 AI* 해당 여부에 대한 영역별 판단 기준과 고영향 AI의 신뢰성 확보 조치의 구체적 이행 방안**을 하위법령에서 규정 ▪ 특히, 고영향 AI는 기본권에 중대한 영향을 미치는 AI시스템으로, ▲사용영역, ▲기본권에 대한 위험의 영향·중대성·빈도, ▲활용 영역별 특수성 등을 종합적으로 고려하여 판단 * 생명·신체·기본권에 중대한 영향을 미치거나 위험 초래 우려가 있는 AI시스템으로 특정 영역(에너지, 보건의료, 원자력, 교통, 교육 등)에서 활용되는 AI시스템 ** ▲위험관리방안(위험관리조직 운영, 위험관리교육 시행), ▲이용자보호방안(보안대책, 오작동 방지대책 수립, 제품·서비스 제공 시 고객 피드백 절차 마련) 등
	▪ **(AI 영향평가)** AI제품·서비스 제공 시 사람의 기본권에 미치는 영향평가*를 실시할 수 있으며, 구체적 내용·방법**을 시행령에서 규정 * 사업자가 자율적으로 실시하되, 고영향 AI에 대해선 영향평가 실시 노력의무 규정 ** 영향 받는 대상 식별, 영향대상 기본권 식별, 영향의 내용·범위, AI 사용행태, 위험의 예방·손실의 복구, 개선 방안 등을 반영하여 평가를 수행하며, 제3자에게 위탁 수행 가능

출처: PwC Analysis

하위법령 초안에 따르면 시행 초기 기업들의 혼란을 최소화할 수 있도록 과태료 계도 기간 운영을 추진 중이며, 이해관계자 의견수렴을 통해 구체적인 계도기간을 확정할 예정입니다. AI 기본법이 유예 없이 2026년 1월부터 일정 기간의 계도를 거쳐 시행하는 것이 가시화됨에 따라 기업들의 AI 기본법 대응에 대한 압박이 더욱 커질 것으로 예상됩니다.

Ⅱ. AI 기본법 대응 방안 및 로드맵

01. AI 기본법, 기업의 도전 과제

AI 기술은 기업의 경쟁력 확보와 서비스 혁신의 핵심 수단으로 자리잡았으며, 이를 활용한 비즈니스 전략은 이제 선택이 아닌 생존의 조건이 되었습니다. 그러나 AI 기본법은 이러한 기술 활용에 있어 법적 책임과 규제 준수를 동시에 요구하며, 기업에게 새로운 형태의 부담과 도전과제를 안기고 있습니다.

AI 기본법은 고영향 AI, 생성형 AI, 고성능 AI에 대한 규제 요건을 제시하며, 기업은 자사 AI 시스템이 해당 범주에 속하는지 판단하고 이에 따른 관리체계를 수립해야 합니다. 문제는 이러한 규제 대응이 단순한 법무 부서의 업무로 끝나지 않는다는 점입니다. AI 기술은 비즈니스, IT, 소비자보호, 정보보호 등 다양한 부서와 기능에 걸쳐 작동하기 때문에, 기업 내부의 이해관계자들은 다양한 도전과제에 직면하게 됩니다.

AI 기본법에서 기인한 내부 이해관계자의 페인 포인트

출처: PwC Analysis

AI 기본법 대응을 위해서 기업은 단순한 기술 대응을 넘어 조직 전반의 전략과 운영체계를 재정비해야 하는 상황에 놓여 있습니다.

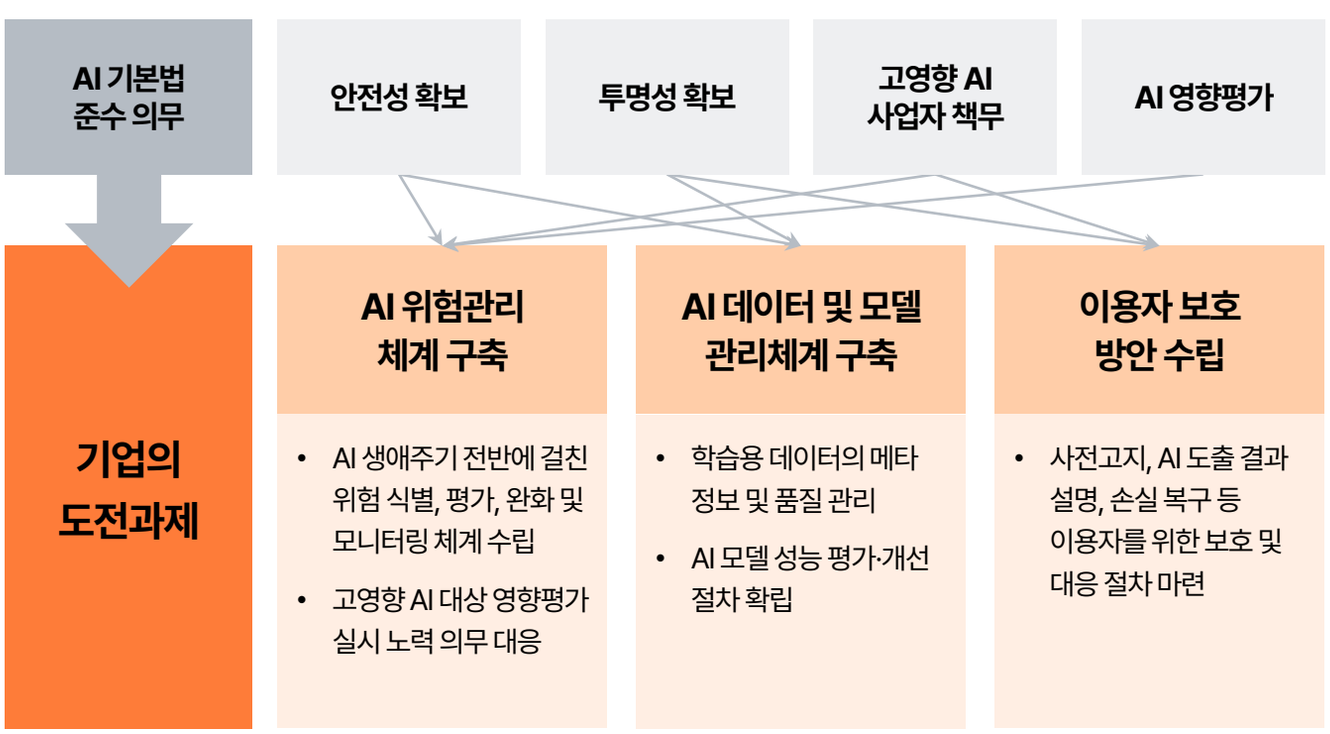
첫 번째 도전과제는 'AI 위험 관리 체계 구축의 복잡성' 입니다. AI 시스템은 설계, 개발, 학습, 운영, 폐기까지 다양한 단계에서 각각 다른 위험을 내포하고 있으며, 이를 식별하고 평가한 뒤 적절히 완화하고 지속적으로 모니터링하는 체계를 수립하는 것은 기업에게 상당한 부담으로 작용합니다. 특히 기존 IT 리스크 관리 체계와의 연계가 불명확한 경우, 중복 관리 또는 관리 공백이 발생할 수 있으며, 실무자들은 AI 위험관리 항목에 대한 이해 부족으로 인해 대응에 어려움을 겪게 됩니다.

두 번째 도전과제는 'AI 학습데이터 및 모델 관리 체계의 부재'입니다. AI 안전성과 투명성 확보를 위해 학습용 데이터의 품질, 출처, 메타 정보 관리와 함께 모델의 성능 평가 및 개선 절차가 요구되고 있지만, 많은 기업은 아직까지 데이터 수집과 모델 운영을 기술 중심적으로 접근하고 있어, 이로 인해 모델의 편향성, 불투명성, 성능 저하 등의 문제가 법적 리스크로 이어질 수 있습니다.

세 번째 도전과제는 '이용자 보호 방안 수립의 실효성 확보'입니다. AI 기본법은 이용자에게 사전고지, 결과 설명, 손실 복구 등의 보호 조치를 제공할 것을 요구하지만, 실제 기업 현장에서는 이러한 조치가 형식적으로 운영되거나 기술적으로 구현되지 않는 경우가 많습니다. 이는 기업의 신뢰도 저하와 함께 법적 책임 발생 가능성을 높이는 요인이 됩니다.

상기 도전과제들은 단순한 규제 대응을 넘어, 기업의 전략, 조직, 기술, 운영 전반에 걸친 구조적 변화와 대응을 요구합니다. 따라서 기업은 AI 기본법 대응에 있어 단기적 관점이 아닌, 장기적 경쟁력 확보를 위해 체계적이고 실용적인 대응 방안을 마련해야 합니다.

AI 기본법 관련 기업의 도전 과제



출처: PwC Analysis

02.
AI 기본법 대응, 거버넌스 구축

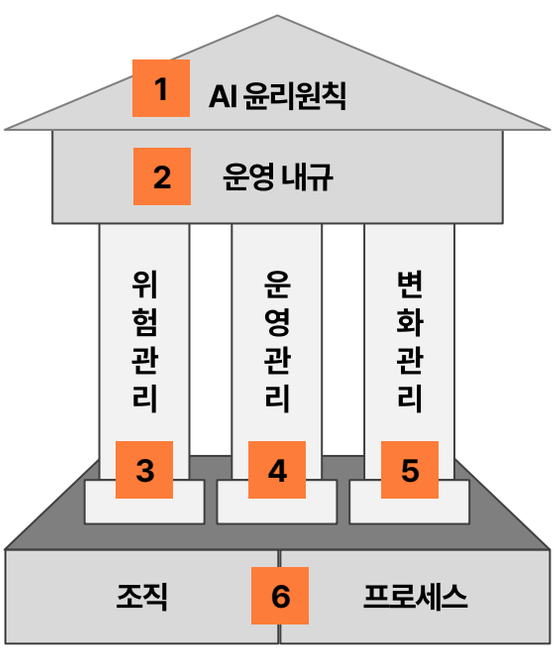
AI 거버넌스 필요성

앞서 설명한 도전과제를 해결하기 위해 많은 기업들이 'AI 거버넌스(AI Governance)' 체계를 도입하고 있으며, 이는 과거 IT 거버넌스나 데이터 거버넌스와 유사한 흐름으로 이해할 수 있습니다.

AI 거버넌스는 단순한 기술 통제 수단이 아니라, 기업의 전략, 운영, 컴플라이언스, 리스크 관리가 통합된 전사적 관리 프레임워크입니다.

AI 거버넌스의 도입은 단순한 규제 대응을 넘어, 기업의 AI 관리 역량을 상향 평준화하고, 조직 간 협업을 촉진하는 효과를 기대할 수 있습니다. 이는 AI 기술을 안전하고 책임 있게 활용함으로써 고객 신뢰를 확보하고, 규제에 안정적으로 대응하며, 장기적으로는 기업의 AI 기반 지속가능한 성장 동력을 확보하는데 기여합니다.

AI 거버넌스 프레임워크



1 **기업의 AI 관련 활동 시 준수할 기본원칙 검토**
(윤리 및 기술 관점 포함)

2 **기업 고유의 AI 거버넌스 관련 업무 기준 및 절차를 고려한 규정 및 매뉴얼 제·개정**

3 **4** **5** **AI 거버넌스 운영 요소**

- 위험관리: 관리대상 식별, 위험수준 평가, 위험완화 조치
- 운영관리: AI 민원대응, 성능관리 등 AI 특화 요소 대응
- 변화관리: AI 거버넌스의 확산 및 내재화

6 **효율적인 운영을 위한 조직 구성, 역할 정의 및 업무 절차 수립**

출처: PwC Analysis

AI 윤리원칙

AI 윤리원칙은 AI 사업자가 AI 서비스의 전 생애주기동안 반드시 지키겠다고 스스로 다짐한 약속입니다.

AI 거버넌스 체계의 핵심 출발점은 바로 'AI 윤리원칙' 수립입니다. 이는 기업이 AI 기술을 책임 있게 활용하기 위한 기본 방향성을 제시하는 기준으로, 모든 임직원이 인지하고 실천해야 할 행동 지침이기도 합니다.

AI 윤리원칙은 단순한 선언적 문구가 아니라, 실제 업무와 의사결정 과정에 내재화되어야 하며, 기술 개발자 뿐 아니라 비즈니스 담당자, 법무, 정보보호, 소비자보호 등 다양한 부서의 이해관계자가 모두 이해하고 준수할 수 있도록 설계되어야 합니다.

AI 윤리원칙을 수립하기 위해서는 먼저 국내외 주요 AI 가이드라인에서 제시하는 핵심 원칙들을 조사·분석하고, 기업 고유의 비전과 산업 특성, 고객 가치, 조직 문화 등을 반영하여 자사에 적합한 원칙으로 재구성해야 합니다.

국내·외 주요 기관의 AI 윤리원칙

과학기술정보통신부

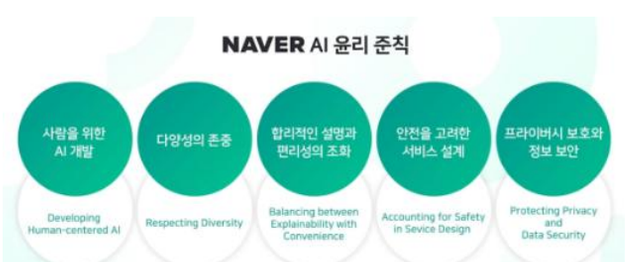
인공지능 윤리기준 : 3대 기본원칙, 10대 핵심요건					
기본 원칙	① 인간의 존엄성 원칙				
	② 사회의 공공선 원칙				
	③ 기술의 합목적성 원칙				
핵심 요건	인권 보장	프라이버시 보호	다양성 존중	침해 금지	공공성
	연대성	데이터 관리	책임성	안전성	투명성

OECD

Values-based AI Principles
① 포용적 성장, 지속가능한 발전 및 복지
② 공정성과 프라이버시를 포함한 인권 및 민주적 가치
③ 투명성 및 설명 가능성
④ 견고성, 보안 및 안전성
⑤ 책임성

출처: 과학기술정보통신부, OECD

국내 기업의 AI 윤리원칙



출처: 각 사 홈페이지

운영 내규	**AI 서비스 개발 · 운영 시에 걸림돌이 아닌, 관련 법규를 준수하기 위한 안전 장치로서의 정책 및 내규 수립이 요구됩니다.**

AI 거버넌스 체계가 실질적으로 작동하기 위해서는 이를 뒷받침하는 운영 내규의 수립이 필수입니다. AI 거버넌스 운영 내규에는 기업의 실제 업무와 의사결정 과정에 적용 가능한 실무 지침과 절차를 포함해야 하며, 이를 통해 조직 내 AI 관련 활동의 일관성과 책임성을 확보할 수 있습니다.

AI 거버넌스 운영 내규는 기존의 규정 체계와의 정합성을 기반으로 제정되어야 하며, 기업마다 이미 정보보안, 개인정보보호, IT 운영, 리스크 관리 등 다양한 분야에서 규정과 지침이 존재하므로, 신설되는 내규는 이러한 기존 체계를 무시하거나 중복하지 않도록 주의해야 합니다. 따라서 내규 수립 시에는 기존 규정과의 중복성 및 상호 연관성을 면밀히 검토하고, AI 관련 활동이 기존 규정에 어떤 영향을 미치는지를 충분히 분석하여, 불필요한 기존 규정 개정은 최소화해야 합니다.

위험관리	**기존 내부통제체계 내에서 관리되는 위험(예: 정보보호, 소비자보호, 보안 등) 뿐만 아니라 AI 특성으로 인한 추가 또는 신규 속성에 대한 위험(예: 편향, 통제력상실, 설명가능성부족 등) 관리가 요구됩니다.**

AI의 특성에 기인한 위험 유형

AI는 기존 IT 시스템과는 달리 자율성, 비가시성, 학습 기반의 의사결정 구조를 갖고 있어, 고유의 위험 유형을 동반합니다. 특히 결과 해석이 어려운 블랙박스 모델, 예측 불가능한 출력 결과, 데이터 편향에 따른 차별 가능성 등 새로운 위험은 기업 신뢰에 직접적인 영향을 줄 수 있습니다. 따라서 기업은 AI 위험관리 체계를 수립하기에 앞서, AI 윤리원칙을 기준으로 삼아 AI 기술 특성에 기인하는 위험유형을 면밀히 식별하고 정의해야 합니다.

AI 위험 유형 예시

성능	AI가 기대된 성능 기준을 만족시키지 못하거나, 의도된 목적에 따라 적절한 결과를 산출하지 못할 위험
공정성	특정 개인이나 집단에게 불공평하거나 편향된 결과를 제공할 위험
설명가능성	AI가 생성한 결과에 대해 기술적 또는 논리적인 관점에서 그 이유나 핵심 요소를 충분히 설명하지 못할 위험
지식재산권	AI에 입력된 정보나 AI가 생성한 결과물이 타인의 지식재산권을 침해할 위험
유해성	AI가 고의적이거나 우발적 상황으로 불법적이거나 사회적·윤리적으로 유해하게 사용될 수 있는 위험
통제가능성	인간이 AI 시스템을 적절히 관리하거나 감독, 제어하지 못할 위험

출처: PwC Analysis

AI 위험관리체계

AI 위험관리 대상으로 식별이 되면 위험요인별로 평가가 이루어지며, 해당 위험수준별로 차등하여 리스크를 점검하고 완화하기 위한 조치를 취하게 됩니다. 최종적으로는 리스크 관리 결과에 대하여 종합적인 검토를 통하여 AI의 위험수준을 관리합니다.

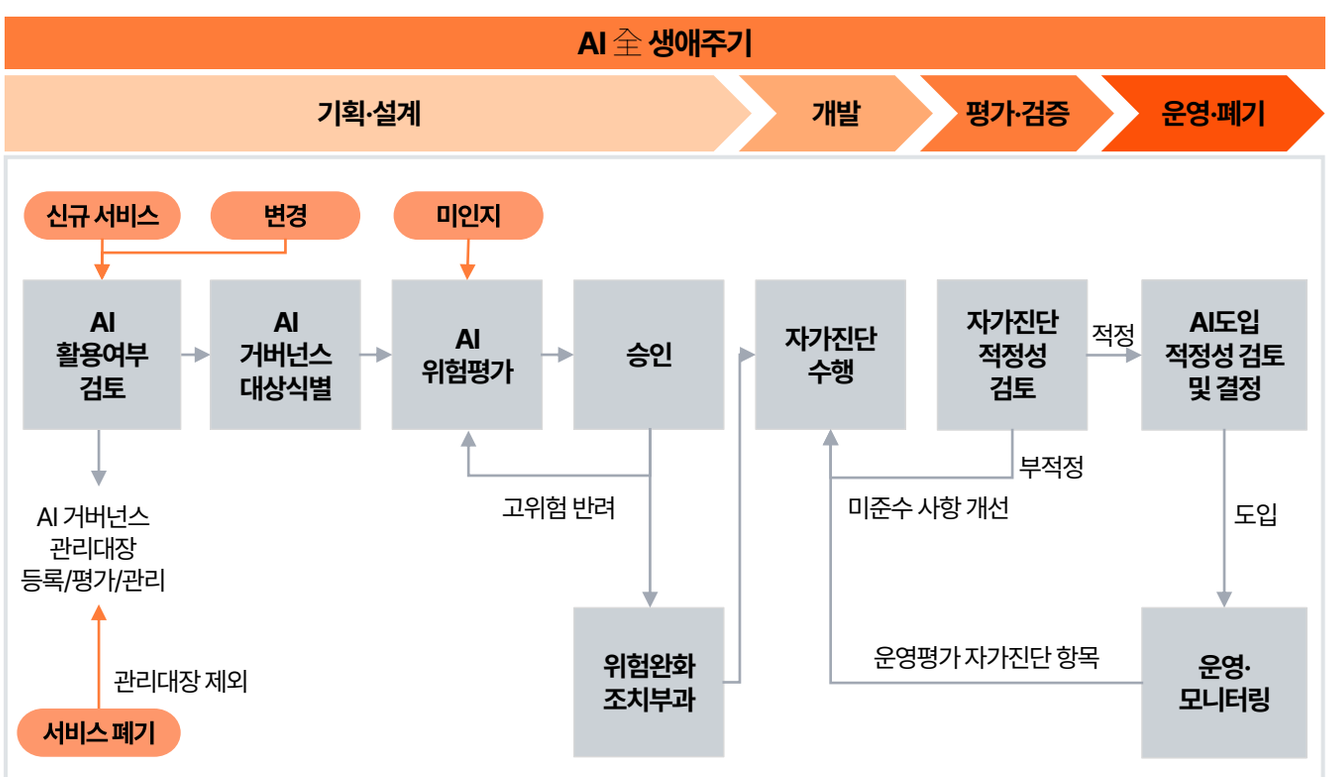
AI 위험관리 단계

1 대상 식별	2 위험 평가	3 리스크 점검·조치	4 심의·결정
AI 리스크 관리 대상 식별	**위험 요인 식별 및 수준 평가**	**체크리스트 준수 여부 진단**	**종합 검토 후 최종 의사결정**
• AI 거버넌스 적용 대상을 명확히 식별하고 선택하는 절차와 기준 설정	• AI 리스크 요소 및 평가항목을 통해 AI 서비스가 본질적으로 갖는 위험의 수준 진단 및 위험등급 산출	• AI 리스크 관리를 위한 단계별 수행 사항, 기술적 장치를 위험등급에 따라 체크리스트 형태로 차등 적용 및 점검	• 리스크 관리 결과에 대해 종합적으로 검토하여 서비스 도입 적정성 여부 최종 검토 및 승인

AI 위험요소 식별 및 위험수준 (고, 중, 저) 평가 → **AI 생애주기별, 위험수준별 점검을 통한 오남용, 부작용 방지**

출처: PwC Analysis

AI 서비스 생애주기 별 위험관리 체계



출처: PwC Analysis

운영관리 **AI 서비스는 기존 IT 서비스이나 시스템보다 한 차원 높은 고객 민원 및 모델 성능관리가 요구됩니다.**

AI 민원 대응

AI 서비스를 운영하는 기업은 이용자로부터 발생할 수 있는 다양한 민원 유형에 대해 사전 대응 방안을 마련해야 합니다.

출력 결과에 대한 설명 요구, 차별 발생, 지식재산권 침해 등 AI 특유의 이슈는 단순한 고객 응대를 넘어, 기술적·법적·윤리적 대응이 요구됩니다.

특히 설명 요구 민원에 대해서는 쉬운 용어 사용, 시각화 된 도표 제공, 결정 과정의 투명성 확보 등 이용자가 이해할 수 있는 방식으로 대응해야 하며, 이는 AI 거버넌스 체계 내에 명확한 절차로 포함되어야 합니다.

설명 요구 민원 대응 예시

이해하기 쉬운 용어 사용

전문 기술 용어 대신 일상적인 용어를 사용하여 설명

단계별 설명 제공

입력 데이터의 처리부터 최종 결과의 도출까지 AI의 작동과정을 순서대로 설명

시각 자료 활용

다이어그램, 차트, 이미지 등 다양한 시각 자료를 활용하여 설명

결정 과정의 투명성 강조

검토 과정을 포함하여 어떤 데이터 및 논리를 통해 도출된 결과인지 명확히 설명

추가 질의/답변 경로 안내

민원인이 충분히 납득하고 이슈가 해소될 수 있도록 추가 질의 및 답변 경로 안내

출처: PwC Analysis

AI 모델 성능평가 지표 예시

기술적 측면	정확성 (Accuracy)	정밀도 (Precision)	재현율 (Recall)
	F1-Score	AUC-ROC curve	RMSE
	BLEU	CIDEr	…
윤리적 측면	SHAP	LIME	Fairness Metrics
	Feature Importance	…	

출처: PwC Analysis

AI 성능관리

AI 서비스의 안정성과 신뢰성을 확보하기 위해서는 기술적 성능 뿐 아니라 윤리적 성능까지 관리해야 합니다.

정확도(Accuracy), 정밀도(Precision) 등 기술적 품질과 공정성, 설명가능성(XAI) 등 윤리적 기준 포괄하는 성능 관리 체계를 마련해야 합니다.

AI 성능 관리는 AI 위험관리에 중요 요소로서, 성능관리를 통해 많은 AI 위험유형을 사전에 식별하고 완화할 수 있습니다.

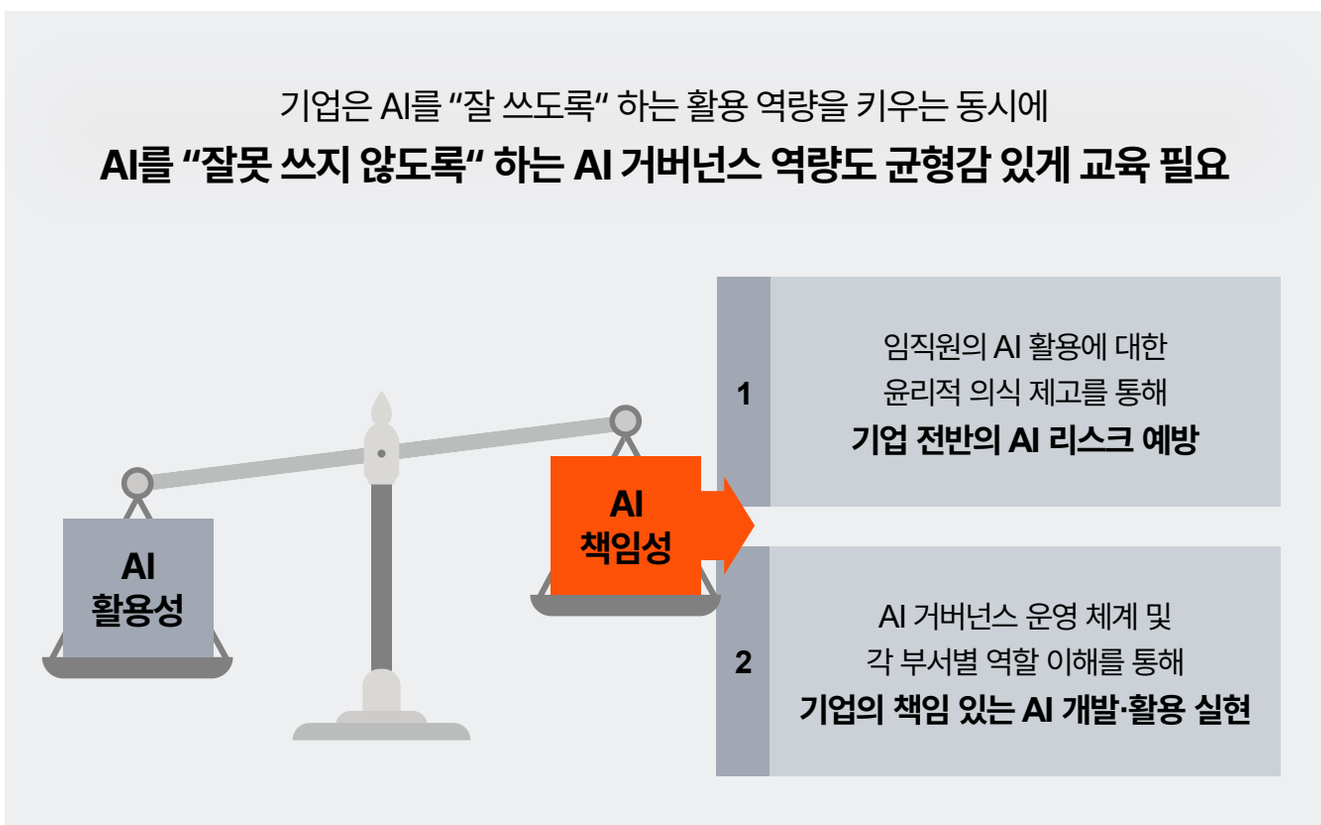
변화관리

AI 리터러시뿐만 아니라 AI 거버넌스 리터러시를 통해 AI 사업자의 모든 임직원이 AI를 잘못 쓰지 않도록 교육과 변화관리가 요구됩니다.

기업이 새롭게 도입하는 AI 거버넌스 체계는 단순한 규정이나 문서화 수준을 넘어, 조직의 AI 활동에 실질적으로 적용될 수 있도록 AI 윤리 교육과 변화관리 활동이 동반되어야 합니다.

AI 윤리 교육은 임직원이 AI의 사회적 영향과 윤리적 책임을 이해하고, 개발·운영 과정에서 이를 실천할 수 있도록 해야 합니다. 기업의 AI 활동에 대한 윤리적 기준을 내재화하기 위해서는 전사적 관점의 교육 프로그램을 마련해야 합니다.

뿐만 아니라, AI 활용에 참여하는 모든 조직 구성원이 AI 거버넌스 체계의 목적과 운영 방식, 역할과 책임을 충분히 이해하고, 실무에 반영할 수 있도록 지속적인 변화관리가 필수입니다. 초기에는 체계에 대한 저항이나 혼란이 발생할 수 있으므로, 단계적 도입과 교육, 피드백 기반의 개선 프로세스를 통해 조직 내 수용성과 실행력을 높여야 합니다.



조직

복잡한 AI 위험을 효과적으로 통제하기 위하여 AI 거버넌스 협의체와 AI 거버넌스 주관부서 중심의 운영 조직체계가 필요합니다.

AI 거버넌스 조직

AI 거버넌스와 같은 신규 업무를 위해서는 특히 조직의 역할과 책임을 명확하게 정의해야 합니다. AI는 다양한 부서와 기능이 연계되어 운영되기 때문에, 단일 부서 중심의 관리로는 복잡한 위험을 효과적으로 통제하기 어렵습니다. 이에 따라 기업은 AI 거버넌스 협의체와 AI 거버넌스 주관부서를 중심으로 한 운영 조직을 구성해야 합니다.

AI 거버넌스 협의체는 AI 관련 주요 의사결정을 수행하는 의사결정기구로서, 다양한 부서의 임원 또는 부서장이 참여하는 분권형 구조로 운영할 것을 권고 합니다. 이를 통해 AI 활동과 관련된 사안을 다각도로 검토하고, 조직 전체의 관점에서 균형 잡힌 결정을 내릴 수 있습니다.

한편, AI 거버넌스 주관부서는 협의체의 결정사항을 실행하고, AI 거버넌스 체계를 실질적으로 운영하는 전담 조직입니다. 이 부서는 AI 정책 수립 및 개선, 위험관리 체계 운영, 교육 및 변화관리 등의 업무를 수행하며, 기업의 AI 활용이 법적·윤리적 기준에 부합하도록 지속적으로 관리할 책임이 있습니다.

조직 구성 시에는 해당 기업의 AI 성숙도를 고려해야 합니다. 이미 AI 전담부서나 전문인력을 충분히 보유한 기업과 AI 관련 조직을 최근에 신설한 기업 간의 AI 역량 차이를 고려하여 유연하게 조직을 설계해야 합니다.

AI 거버넌스 조직도 및 주요 구성요소

AI 거버넌스 협의체

- **AI 유관부서 의견 조율, AI 거버넌스 관련 최종 의사결정 조직**
  - AI 거버넌스 정책, 내규 제·개정부터 전담 조직 통제까지 여러 기능 수행

AI 거버넌스 주관부서

- **AI 거버넌스의 컨트롤 타워**
  - AI 서비스 현황 파악부터 AI 위험관리 지침 배포 및 AI 서비스 검증까지 다양한 스펙트럼의 업무 수행

출처: PwC Analysis

03.
AI 기본법 대응 로드맵

AI 기본법 시행이 다가옴에 따라 기업은 선제적으로 대응 체계를 마련해야 합니다. AI 기본법 대응 체계 구축에 있어 기업은 단기적인 규제 대응뿐만 아니라 중장기적으로 AI의 경쟁력 강화까지 고려하여 전사 차원의 명확한 AI 거버넌스 목표 및 추진 방향성을 갖고 계획을 수립할 필요가 있습니다. 이에 따라 AI 기본법에 효과적으로 대응하기 위한 3단계 로드맵을 제안합니다.

1단계

AI 기본법에 대한 Readiness 진단 및 대응 전략 수립

기업은 1) 현재 운영 중인 내규와 조직 구성에 대한 AI 기본법 준수 수준, 2) AI 개발 및 활용 전체 생애주기상 사업자의 의무를 준수하고 있는지에 대한 진단이 필요합니다. 이를 기반으로 AI 거버넌스의 추진 목적 및 방향성 (위험 통제 및 예방 vs. 혁신 추진)과 AI 거버넌스 관리 대상을 정의할 수 있습니다.

2단계

AI 거버넌스 수립

기업의 현황과 AI 기본법에 대한 대응 전략에 따라 기업 맞춤형 AI 거버넌스 체계를 수립합니다. 기업의 AI 윤리 원칙을 수립하고 기업의 고유 업무 기준과 절차를 고려한 AI 규정, 매뉴얼을 제·개정합니다. 또한 위험관리, 운영관리, 변화 관리 방안을 정의합니다. 이를 효율적으로 운영하기 위한 조직 구성, 역할 및 업무 절차를 마련합니다.

3단계

효율적 AI 거버넌스 실행을 위한 IT 시스템 구축

전사의 모든 서비스에 AI가 도입되고, AI 서비스의 자율화 및 고도화 수준이 높아질수록 AI 서비스에 대한 위험은 점차 높아지게 됩니다. 이에 따라 AI 기본법에서 요구하는 위험관리 수행 및 이력 추적 관리, 고객에 대한 민원 대응 등에 대한 수작업에는 한계가 있습니다. AI 거버넌스 체계에 대한 시스템화를 통하여 보다 투명하고 효율적인 체계 수립이 가능해집니다.

AI 기본법 대응 로드맵

1단계	2단계	3단계
AI 기본법에 대한 Readiness 진단 및 대응 전략 수립	**AI 거버넌스 수립**	**효율적 AI 거버넌스 실행을 위한 IT 시스템 구축**
▪ 고영향 AI 여부 검토 ▪ 법적 의무 이행 수준 점검 ▪ 거버넌스 방향성과 관리 기준 설정	▪ AI 윤리 원칙 및 내부 규정 정비 ▪ 위험·운영·변화 관리 방안 마련 ▪ 조직 구성 및 역할 정의	▪ 위험관리 및 이력관리 기능 구현 ▪ 생성형 AI 고지 및 민원 대응 (설명가능성) 시스템 마련 등 법률상 요구되는 기능 지원

출처: PwC Analysis

04.
결언 및 시사점

AI 기본법은 기업에게 단순한 규제가 아닌, AI 기술을 책임 있게 활용하기 위한 제도적 기준을 제시합니다. 특히 고영향·생성형 AI를 개발·제조·제공하는 기업은 투명성, 안전성 등 다양한 법적 의무를 이행해야 하며, 이를 위해 내부 거버넌스를 갖춰야 합니다. 하위법령과 가이드라인이 구체화되고 AI 기본법이 시행되면 기업의 대응 부담은 커질 수 있으나, 이는 동시에 신뢰 기반의 AI 활용을 통해 지속가능한 경쟁력을 확보할 수 있는 기회이기도 합니다.

AI 기본법 대응은 기업에 복잡한 과제를 안깁니다. 첫째, 자사의 AI가 고영향·고성능·생성형 AI에 해당하는지 판단하고, 이에 맞는 위험관리 체계를 구축해야 합니다. 둘째, AI의 안전성과 투명성을 확보하기 위해 학습용 데이터 및 모델 관리 체계를 개선해야 합니다. 셋째, 실효성 있는 이용자 보호 방안을 수립해야 합니다. 이 모든 과제는 법무뿐 아니라 조직 전체의 전략과 운영 체계 재정비를 요구하며, 기업은 장기적 관점에서 대응해야 합니다.

이러한 복합적인 문제에 대응하기 위해서는 기업의 전략, 운영, 컴플라이언스, 리스크 관리가 통합된 전사적 관리 프레임워크로서 AI 거버넌스가 필요합니다. AI 서비스는 다양한 부서와 밸류체인에 걸쳐 운영되며, 자율성과 적응성을 갖춰 예측이 어려운 위험을 동반합니다. 따라서 단일 부서 중심의 대응으로는 복잡한 위험을 효과적으로 통제하기 어렵습니다.

AI 거버넌스는 AI 기본법이라는 외부 규제에 대응하기 위한 수단이자, 내부적으로는 AI 기술을 전략적으로 활용하고 통제할 수 있는 기업의 핵심 역량으로 자리잡아야 합니다.

AI 기본법 시행을 앞둔 지금이야말로 기업이 AI 거버넌스를 통해 혁신 추진과 안전의 균형을 이루고, AI 혁신의 기초체력을 갖추어 AX 시대에 앞서가는 경쟁력을 확보할 수 있는 최적의 시점입니다.

PwC AI Trust Center 소개

PwC AI Trust Center의 서비스

PwC AI Trust Center는 기업의 AI 기본법 대응부터 인증 지원, Data 및 IT 거버넌스, 성능, 보안까지 안전하고 신속하게 AI 경쟁력을 갖출 수 있도록 종합 서비스를 제공합니다.

외부 규제 및 가이드 준수

AI 기본법 대응 체계 구축

AI 전 생애주기에 안전성, 신뢰성을 확보하고 이용자를 보호하도록 원칙, 내규, 조직, 프로세스 정비

AI 관련 국제 표준 인증(ISO) 지원

인증·심사에 필요한 문서 정리, 보완과 모의 점검, 임직원 교육 전 과정 지원

AI 보안

최신 국내외 보안 규정·가이드 기준 데이터, 모델, 운영 보안 정책-절차-점검표-훈련 시나리오 수립, 적용

종합 컨설팅 서비스 제공

AI를 위한 IT 거버넌스

기존 IT의 거버넌스 내에서 AI를 효율적으로 관리할 수 있도록 정비하고 AI 거버넌스와 연계 지원

AI 성능 평가

AI 서비스 유형별 기술적 성능 (정확성 등)과 윤리적 성능 (편향 등) 평가를 위한 방법론, 지표 제공

AI를 위한 데이터 거버넌스

AI 개발 속도와 성능을 극대화할 수 있도록 AI가 이해할 수 있는 형태로 데이터 생성·관리·최신화 체계 마련

AI 개발·운영 관리 체계 고도화

AI Trust 기반 기업의 변화

AI Trust를 내재화한 기업은 흔들리지 않는 AI 혁신의 기초체력을 갖춤으로써, AX 시대에 앞서가는 경쟁력을 확보하게 될 것입니다.

01. 신규 규제 대응 및 Compliance 강화

외부 규제는 준수하되, 내부 혁신을 저해하지 않도록 균형 잡힌 AI 기본법 대응체계를 장착

02. 기존 운영체계의 AI-enabled 확장

IT, 데이터, 사이버 보안 등 기존 영역에서 AI로 인해 변화가 필요한 사항을 진단하고 개선 방안을 수립

AI 경쟁력 확보

Contacts

문 홍 기 Partner

hong-ki.moon@pwc.com
02-709-0394

김 은 섭 Partner

eun-seop.kim@pwc.com
02-3781-9749

김 진 유 Partner (센터장)

jin-you.kim@pwc.com
02-3781-1486

김 의 준 Partner

euijun.kim@pwc.com
02-709-8969

박 현 출 Partner

hyunchul.park@pwc.com
02-709-0412

이 성 호 Partner

sungho1.lee@pwc.com
02-3781-1773

정 성 문 Partner

sungmoon.cheong@pwc.com
02-709-4012

박 유 현 Partner

yuhyeon.park@pwc.com
02-709-8811

S/N: 2509C-RP-102