

Artificial Intelligence Law Issue Brief

Transitioning from AI Legislation to Compliance
by Design: An Australian Perspective
AI 법률에서 컴플라이언스 설계로의 전환: 호주의 관점

Niloufer Selvadurai
닐로퍼 셀바두라이

E

U

Transitioning from AI Legislation to Compliance by Design: An Australian Perspective



Niloufer Selvadurai (Macquarie Law School, Professor)*

Introduction

- As nations around the world seek to design AI regulations to mandate the responsible use of artificial intelligence (AI),¹ it is useful to consider the evolving law reform discourse in particular countries.² For while all nations agree on the need to enact legislation to control the use and operation of AI systems and services, they are adopting varying approaches to this challenge. In September 2024, the Australian Government released a paper entitled Safe and

* BA LLB (Hons I) University of Sydney, PhD, Macquarie University; Director of Research & Innovation, Macquarie Law School; Co-Chair, Data Horizons Research Centre <<https://www.mq.edu.au/research/research-centres-groups-and-facilities/centres/data-horizons>> Personal webpage <<https://researchers.mq.edu.au/en/persons/niloufer-selvadurai>>

1) 'AI' is a broad term that encompasses a broad range of computational efforts to replicate human cognitive capacity: A.Y. Bokovnyal et al, "Pressing Issues of Unlawful Application of Artificial Intelligence" (2020) 9 International Journal of Criminology and Sociology 1054-1057.

2) See further – Mathijis Maas, 'Aligning AI Regulation to Sociotechnical Change' in Justin Bullock et al (eds), *The Oxford Handbook of AI Governance* (Oxford University Press, 2022) 361-362; Luciano Floridi, et al, 'An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations,' (2018) 28(4) *Minds & Machines* 689; Iason Gabriel, 'Artificial Intelligence, Values, and Alignment' (2020) 30(3) *Minds & Machines*. 411; Luciano Floridi & Josh Cows, 'A Unified Framework of Five Principles for AI in Society' (2019) 1(1) *Harv. Data Science Review* <<https://hdsr.mitpress.mit.edu/pub/10jsh9d1/release/8>> Accessed 10 November 2024; As to broader societal challenges see book Nikos Kourtas & Niloufer Selvadurai, *Recreating Creativity, Re-inventing Inventiveness: AI and IP* (2024, Routledge UK).

Responsible AI in Australia: Proposals Paper for Introducing Mandatory Guardrails for AI in High-risk Settings.³⁾ This builds on the the August 2024 Voluntary AI Standard⁴⁾ and the January 2024 Safe and Responsible AI in Australia.⁵⁾ However, there are certain issues of uncertainty as to the likely design and operation of any AI legislation. In the context of the opaque and ubiquitous nature of AI use, these include challenges relating to the enforcement of such legislation arising from the difficulty of identifying offending behaviour, identifying liable entities and gathering evidence to support prosecution. It is here suggested that an additional dimension to the design of AI legislation is embedding technical controls within AI algorithms. I have previously described this innovative idea and have called it ‘compliance by design.’ In such a context, the objective of this paper is to evaluate the Australian AI law reform discourse and consider the need to support legislation with such technical controls and safeguards.

3) Australian Government, Department of Industry, Science and Resources, Safe and Responsible AI in Australia Proposals Paper for Introducing Mandatory Guardrails for AI in High-risk Settings, September 2024 <[https://www. file:///C:/Users/nilou/Desktop/2024%20AI%20for%20Korean%20Legal%20 Research %20Institute/proposals_paper_for_introducing_mandatory_guardrails_for_ai_in_high_risk_settings.pdf](https://www.file:///C:/Users/nilou/Desktop/2024%20AI%20for%20Korean%20Legal%20Research%20Institute/proposals_paper_for_introducing_mandatory_guardrails_for_ai_in_high_risk_settings.pdf)> Accessed 12 November 2024. Hereafter referred to as the ‘Proposals Paper for Introducing Mandatory Guardrails for AI in High-risk Settings’.

4) Australian Government, Department of Industry, Science and Resources, Voluntary AI Safety Standard, August 2024 <<https://www.industry.gov.au/publications/voluntary-ai-safety-standard>> Accessed 8 November 2024.

5) Australian Government, Department of Industry, Science and Resources, The Australian Government’s Interim Response to Safe and Responsible AI Consultation, January 2024 <<https://www.industry.gov.au/news/australian-governments-interim-response-safe-and-responsible-ai-consultation>> Accessed 8 November 2024.

1. An International Overview

- In 2024, nations around the world have released a variety of law reform papers and announcements as to AI legislation. In March 2024, the European Union approved the historic EU AI Act.⁶⁾ In July 2024, the UK Government announced its intention to introduce an AI Bill.⁷⁾ So far, six states of the United States (US) have enacted legislation to govern specific aspects of AI, such as the use of AI-generated automated systems or the use of AI in education.⁸⁾ Canada has released a proposal on AI and Data Act.⁹⁾ In March 2024, the United Nations General Assembly adopted a landmark resolution on the promotion of “safe, secure and trustworthy” artificial intelligence (AI) systems seeks to support sustainable development.¹⁰⁾

2. The 2024 Australian Proposal for Mandatory AI Guardrails

2.1 Objectives

- The objective of the September 2024 Proposals Paper for Introducing Mandatory Guardrails for AI in High-risk Settings is to consider options for the design of mandatory legislative guardrails for the use of AI in high-risk settings. The overarching purpose is to build a better future for all Australians through enabling a productive, resilient and sustainable economy, enriched by science and technology. An outline of the objectives of the paper are represented in Figure 1. The proposed legislation seeks to expand and mandate the 2019 Australian AI Ethics Principles.

6) European Parliament, Artificial Intelligence Act: MEPs adopt landmark law, 13 March 2024.

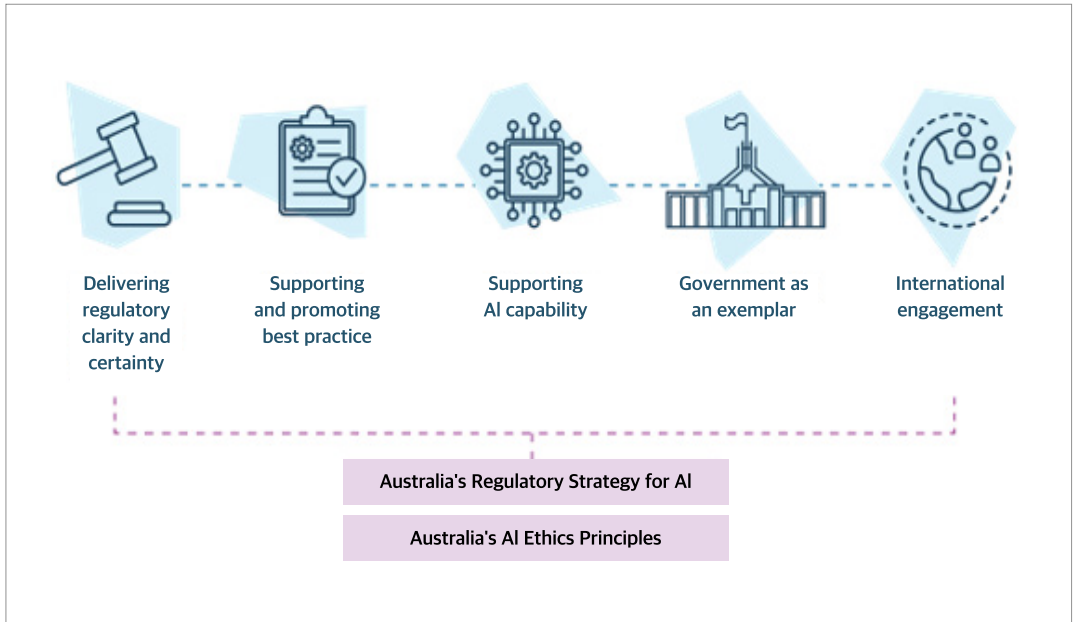
7) United Kingdom, Oral Statement to Parliament, The King’s Speech 2024, 17 July 2024. <https://www.gov.uk/government/speeches/the-kings-speech-2024> Accessed 12 November 2024.

8) US National Conference of State Legislatures, Summary: Artificial Intelligence 2024 Legislation, USA Government, 3 June 2024.

9) Parliament of Canada, Minister of Innovation, Science and Industry, Amendments to Bill C-27 (44-1), 28 November 2023.

10) United Nations, General Assembly Adopts Landmark Resolution on Artificial Intelligence, UN News, 21 March 2024. <https://news.un.org/en/story/2024/03/1147831#:~:text=21%20March%202024%20SDGs%20The%20UN%20General%20Assembly,that%20will%20also%20benefit%20sustainable%20development%20for%20all.>

<Figure 1> The Objectives of Australian AI Legislation

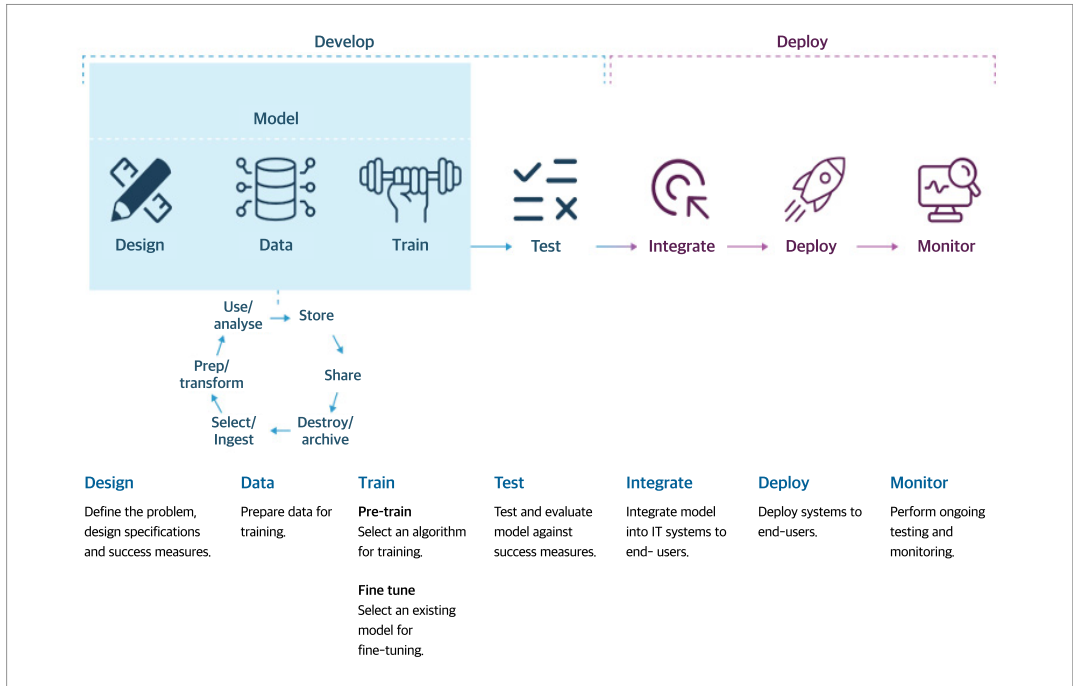


Source: Australian Government, Department of Industry, Science and Resources, Safe and Responsible AI in Australia Proposals Paper for Introducing Mandatory Guardrails for AI in High-risk Settings, September 2024 at p. 6.

2.2 Proposed Ambit of Operation

- The proposed guardrails in the 2024 Proposals Paper for Introducing Mandatory Guardrails for AI in High-risk Settings will apply across the full lifecycle of AI design, testing and deployment. This AI lifecycle is represented in Figure 2 below.

<Figure 2> The AI Lifecycle

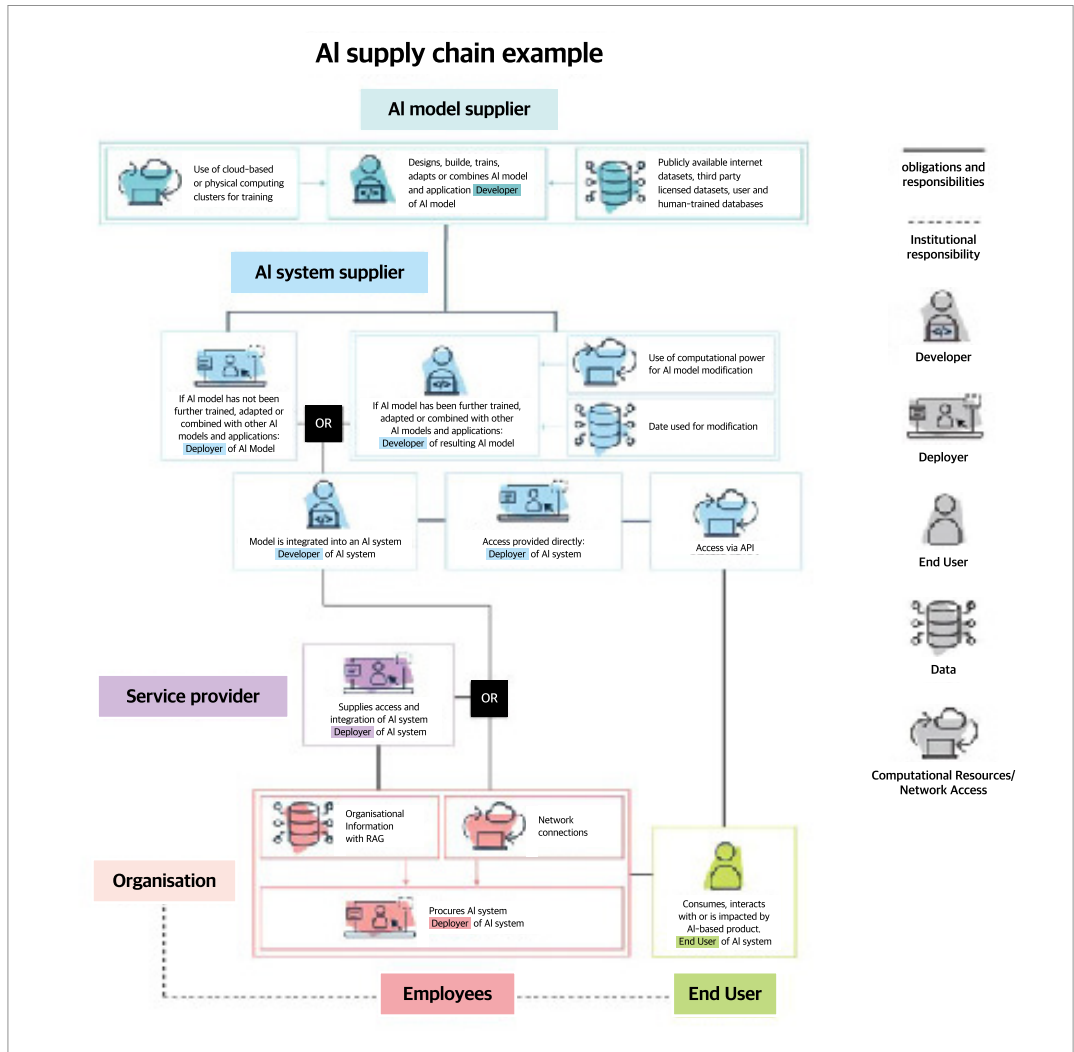


Source: Australian Government, Department of Industry, Science and Resources, Safe and Responsible AI in Australia Proposals Paper for Introducing Mandatory Guardrails for AI in High-risk Settings, September 2024 at p. 10.

- Additionally, the mandated guardrails will apply across the AI supply chain, from the AI model supplier to the AI system supplier to the service provider to the organisation to the employee to the end user. An example of how proposed legislation will apply is represented in Figure 3 below.

2.3 Regulating the AI Supply Chain

<Figure 3> The AI Supply Chain



Source: Australian Government, Department of Industry, Science and Resources, Safe and Responsible AI in Australia Proposals Paper for Introducing Mandatory Guardrails for AI in High-risk Settings, September 2024 at p. 34.

2.4 The Australian AI Strategy

- Finally, the paper usefully summarises the key elements in the proposed holistic AI regulatory strategy. Figure 4 below outlines the key elements in the Australian government’s AI regulatory strategy.

<Figure 4> The Australian Approach to AI Regulation

Key elements of whole-of-government AI regulatory strategy	
Setting clear regulatory objectives and principles	<p>Such as:</p> <ul style="list-style-type: none"> • Protect people from harm • Innovation for economic benefit • Societal wellbeing
Clarifying the regulatory approach	<p>Such as:</p> <ul style="list-style-type: none"> • The application of a risk-based approach • Defining harms, specifying the types and levels of AI risk • Grounding in human rights
Ensuring the regulatory approach is fit-for-purpose for Australia’s unique settings	<p>Having regard to:</p> <ul style="list-style-type: none"> • AI as an identified critical technology impacting Australia’s national interest, including economic prosperity, national security and social cohesion • Australia’s economic interests and comparative advantage
Ensuring consistency in regulatory approach across government	<p>Ensuring alignment in approach and clarifying obligations across the public and private sector including:</p> <ul style="list-style-type: none"> • Establishment of mandatory guardrails • Strengthening and clarifying existing laws • Voluntary AI Safety Standard • Government as an exemplar
Promoting clarity and consistency of legal terms and rules	<p>Such as:</p> <ul style="list-style-type: none"> • Improving consistency of legal terms and cross-sectoral clarity of existing laws that apply to the development, deployment and use of AI • Clarifying legal liability • Clarifying interaction with state and territory laws

Key elements of whole-of-government AI regulatory strategy	
Clarifying the application of existing and new laws	<ul style="list-style-type: none"> • Clarify and strengthen existing laws (such as privacy, consumer protection, IP) • Clarify the interaction between existing laws and any new mandatory guardrails • Clarify the role of sectoral regulators
Clarifying the purpose of hard law and soft law measures	<p>Considering the role of:</p> <ul style="list-style-type: none"> • Hard law in setting minimum expectations for the conduct of developers and deployers of AI systems • Soft law in providing added guidance and supporting transition of industry governance practice
Specifying the approach to stakeholder consultation across regulatory initiatives	<p>Including:</p> <ul style="list-style-type: none"> • Genuine, ongoing consultation with industry, unions, civil society and academia • Inclusive approaches to engaging potentially impacted communities and their representatives across civil society • Ongoing independent expert advice including legal, policy and technical knowledge
Determining the settings necessary to facilitate an effective and responsive regulatory ecosystem	<p>Considering:</p> <ul style="list-style-type: none"> • Regulator capabilities, resources and powers necessary to enforce the law • Practical guidance to ensure legal and ethical responsibilities are understood and upheld • Education and resources to support the exercise of individual rights • Support for SMEs to understand and meet obligations
Having regard to international alignment and interoperability	<p>Ensuring Australia's regulatory approach has regard to:</p> <ul style="list-style-type: none"> • International alignment and interoperability, especially for Australian firms operating across borders • International normative grounding of AI laws in human rights

Source: Australian Government, Department of Industry, Science and Resources, Safe and Responsible AI in Australia Proposals Paper for Introducing Mandatory Guardrails for AI in High-risk Settings, September 2024, Attachment C at p. 59.

3. Issues of Continuing Uncertainty

- There are, however, a variety of areas of uncertainty in the proposed AI mandatory guardrails in high-risk settings. Firstly, there is some uncertainty as to what constitutes a ‘high risk setting’. In this regard, the paper outlines two potential categories of high-risk AI. The first relates to situations where uses of the particular AI system or model is either known or foreseeable. In such cases, risk can be assessed with reference to the factual context. A second category relates to AI systems and models for which the potential risks cannot be reasonably be foreseen. The paper outlines a set of principles that can be used for the purposes of determining whether a particular AI system is high risk. These include ‘the risk of adverse impacts to an individual’s rights recognised in Australian human rights law without justification,’¹¹⁾ ‘the risk of adverse impacts to an individual’s physical or mental health or safety,’¹²⁾ ‘the risk of adverse legal effects, defamation or similarly significant effects on an individual,’¹³⁾ ‘the risk of adverse impacts to groups of individuals or collective rights of cultural groups,’¹⁴⁾ and ‘the risk of adverse impacts to the broader Australian economy, society, environment and rule of law.’¹⁵⁾ However, the Proposals paper does not provide a risk assessment framework for the identification and assessment of such risks. Moreover, principles such as the risk of adverse impacts to the ‘broader Australian economy, society, environment and rule of law’ is very widely framed, and will be uncertain in application, leading to potential inconsistencies.
- Secondly, it is not wholly clear how the proposed mandatory guardrails will align with existing laws. While Australia does not have broad AI specific laws, there are a variety of general laws, that do presently apply. For example, consumer law, such as the prohibition against false or misleading representations apply to AI-generated advice,

11) Australian Government, Department of Industry, Science and Resources, Safe and Responsible AI in Australia Proposals Paper for Introducing Mandatory Guardrails for AI in High-risk Settings, September 2024, p. 19, principle a.

12) Above note, principle b.

13) Above note, principle c.

14) Above note, principle d.

15) Above note, principle e.

as does the established law on negligent misstatement. In charting a way forward, the paper presents options for reform. These include a ‘domain specific approach’ which involves adopting the guardrails within existing regulatory frameworks as needed, a ‘framework approach’ which involves enacting new AI-specific legislation to amend existing legislation or a ‘whole of economy approach’ which involves enacting a new AI-specific Act such as for example an AI Act. The paper does not however, proceed to map existing laws that may apply to AI or seek to delineate their order of operation. Such a step, as undertaken in the United Kingdom in its 2024 Implementing the UK’s AI Regulatory Principles Report would have been a useful progression to the analysis.¹⁶⁾

- Finally, it is not clear what enforcement and administration entities and infrastructure will be used to monitor the legislation and prosecute misfeasance. In light of the non-transparent nature of AI, it is often difficult to discern its operation and impact. Further, given the ubiquitous nature of AI it is also costly and logistically difficult to monitor AI use and prosecute for misfeasance. In such a context, it is suggested that a new approach to AI regulation is needed. I suggest an approach of compliance by design - this will be the focus of the next seton of this paper.

4. Transitioning from Legislation to Compliance by Design

- Compliance by design involves embedding legislative mandates and standards within AI algorithms, tools and controls. This concept was first developed in my article entitled ‘Advancing Lawful AI through Compliance by Design’.¹⁷⁾ The proposed compliance by design framework requires a close interdisciplinary collaboration between technology lawyers and computer scientist who design, test and deploy AI systems. It also requires the involvement of risk assessment experts to assess AI risk, and support the design of regulation that aligns with the

16) United Kingdom, Department for Science Innovation & Technology, Implementing the UK’s AI Regulatory Principles Report (2024).

17) Niloufer Selvadurai, ‘Advancing AI through “Compliance by Design” ’ (2025) 31(1) Computer and Telecommunications Law Review 1; The concept is broadly applied in the context of blockchain smart contract regulation in Niloufer Selvadurai, ‘Mitigating the legal challenges associated with blockchain smart contracts: the potential of hybrid on-chain/off-chain contracts,’ (2023) 80(3) Washington & Lee Law Journal 1163-1179.

risk nature and level of a particular AI system or model. This also involves appropriately calibrating the risk with the potential productivity and efficiency gains of the AI. Finally, it is critical to integrate the perspectives of domain experts, such as those in the public sector, health and finance, who can advise on how a particular AI system will operate within their industry or societal group. It is important to note that such a compliance by design framework is not a substitute for legislation but rather a supplementary framework to help achieve the overarching objectives of the legislation and effectively enforce the law.

Conclusion

- Therefore, as law and policy makers around the world consider how best to regulate AI so as to support its lawful and responsible use, while also supporting its productivity and economic benefits, the Australian Proposals Paper for Introducing Mandatory Guardrails for AI in High-risk Settings provides some useful insights of the matters that need to be addressed when designing future laws. It is submitted that any AI legislation should be further supported by a 'compliance by design' framework to embed compliance within the design of AI systems and models and support effective enforcement of the law.

AI 법률에서 계획적 준수로의 이행: 호주의 관점



Niloufer Selvadurai (맥쿼리대학교 로스쿨 교수)*

개요

- 전 세계 국가가 인공지능(AI)의 책임 있는 사용을 의무화하기 위하여 AI 규정을 계획하고 있으므로,¹⁾ 특정한 국가에서 활발히 진행되고 있는 법률 개혁 담론을 고려하는 것이 좋을 것이다.²⁾ 모든 국가가 AI 시스템과 서비스의 사용 및 운영을 통제하는 법률을 제정할 필요성에 대해서는 동의하고 있지만, 이러한 과제에 대해 다양한 접근법을 채택하고 있기 때문이다. 2024년 9월, “호주 정부는 호주의 안전하고 책임 있는 AI: 고위험

* 시드니 대학교 법학사(우등), 맥쿼리 대학교 박사, 맥쿼리 로스쿨 연구혁신 책임자, 데이터 호라이즌 연구센터 공동의장 <<https://www.mq.edu.au/research/research-centres-groups-and-facilities/centres/data-horizons>> 개인 웹페이지 <https://researchers.mq.edu.au/en/persons/niloufer-selvadurai>

1) ‘AI’는 사람의 인지능력을 모사하려는 광범위한 컴퓨터 사용 활동을 아우르는 광의의 용어이다. A.Y. Bokovnyal et al, “Pressing Issues of Unlawful Application of Artificial Intelligence” (2020) 9 International Journal of Criminology and Sociology 1054-1057.

2) 보다 자세한 내용은 Mathijis Maas, ‘Aligning AI Regulation to Sociotechnical Change’ in Justin Bullock et al (eds), The Oxford Handbook of AI Governance (Oxford University Press, 2022) 361-362; Luciano Floridi, et al, ‘An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations,’(2018) 28(4) Minds & Machines 689; Iason Gabriel, ‘Artificial Intelligence, Values, and Alignment’ (2020) 30(3) Minds & Machines. 411; Luciano Floridi & Josh Cowls, ‘A Unified Framework of Five Principles for AI in Society’ (2019) 1(1) Harv. Data Science Review <<https://hdr.mitpress.mit.edu/pub/10jsh9d1/release/8>> Accessed 10 November 2024; As to broader societal challenges see book Nikos Koutras & Niloufer Selvadurai, Recreating Creativity, Re-inventing Inventiveness: AI and IP (2024, Routledge UK) 참조.

환경에서 AI에 대한 의무적 보호장치를 도입하기 위한 제안”이라는 문서를 배포하였다.³⁾ 이 문서는 2024년 8월 자발적 AI 표준⁴⁾과 2024년 1월 호주의 안전하고 책임 있는 AI를 기반으로 한다.⁵⁾ 다만, AI 입법의 계획과 운용 가능성과 관련하여 불확실성이라는 특정한 문제가 존재한다. AI 사용이 투명하지 않고 보편적인 특성이 있다는 맥락에서 이러한 문제는 범법행위 식별, 책임 주체 식별 및 기소를 뒷받침할 증거 수집의 어려움으로 인하여 발생하는 법률의 집행에 관한 과제가 포함된다. 여기에서는 AI 법률 계획에 대한 추가적인 차원으로 AI 알고리즘 내에 기술 통제를 내포한다는 것을 제안한다. 필자는 이미 이와 같은 혁신적인 사고를 서술하였고, 이를 ‘계획적 준수’라 하였다. 이러한 맥락에서 이 논문의 목적은 호주 AI 법률 개혁 담론을 평가하고 기술 통제 및 안전장치를 갖춘 법률을 지원할 필요성을 고려하는 것이다.

3) Australian Government, Department of Industry, Science and Resources, Safe and Responsible AI in Australia Proposals Paper for Introducing Mandatory Guardrails for AI in High-risk Settings, September 2024 <https://www.file:///C:/Users/nilou/Desktop/2024%20AI%20for%20Korean%20Legal%20Research%20Institute/proposals_paper_for_introducing_mandatory_guardrails_for_ai_in_high_risk_settings.pdf> Accessed 12 November 2024. 이하 ‘고위험 환경에서 AI에 대한 의무적 보호장치를 도입하기 위한 제안’이라 한다.

4) Australian Government, Department of Industry, Science and Resources, Voluntary AI Safety Standard, August 2024 <<https://www.industry.gov.au/publications/voluntary-ai-safety-standard>> Accessed 8 November 2024.

5) Australian Government, Department of Industry, Science and Resources, The Australian Government’s Interim Response to Safe and Responsible AI Consultation, January 2024 <<https://www.industry.gov.au/news/australian-governments-interim-response-safe-and-responsible-ai-consultation>> Accessed 8 November 2024.

1. 국제적 개관

- 2024년, 전 세계 국가는 AI 법률에 관해 다양한 법률 개혁 논문과 담화를 발표하였다. 2024년 3월, 유럽 연합은 역사적인 EU AI 법을 승인하였다.⁶⁾ 2024년 7월, 영국 정부는 AI 법안을 도입할 의사가 있다고 발표했다.⁷⁾ 현재까지 미국(US)의 6개 주가 생성형 AI 자동시스템의 사용 또는 교육분야의 AI 사용 등 AI의 특정한 측면에 적용되는 법률을 제정하였다.⁸⁾ 캐나다는 AI 및 데이터 법에 관한 제안을 발표하였다.⁹⁾ 2024년 3월, 유엔총회는 지속 가능한 개발을 지원하기 위한 “안전하고 신뢰할 수 있는” 인공지능(AI) 시스템의 촉진에 관한 중요한 결의안을 채택하였다.¹⁰⁾

2. 2024년 호주 의무적 AI 보호장치

2.1 목표

- 2024년 9월 고위험 환경에서 AI에 대한 의무적 보호장치를 도입하기 위한 제안의 목표는 고위험 환경에서 AI 사용을 위한 의무적 입법 보호장치의 계획에 대한 선택사항을 고려하는 것이다. 가장 중요한 목적은 과학과 기술을 통해 풍부해진 생산적이고 탄력적이며 지속 가능한 경제를 가능하게 함으로써 모든 호주인을 위한 보다 나은 미래를 구축하는 것이다. 이 논문의 목표에 관한 개요는 그림 1에서 설명한다. 제안한 법률은 2019년 호주 AI 윤리 원칙을 확대하고 의무화하려는 것이다.

6) European Parliament, Artificial Intelligence Act: MEPs adopt landmark law, 13 March 2024.

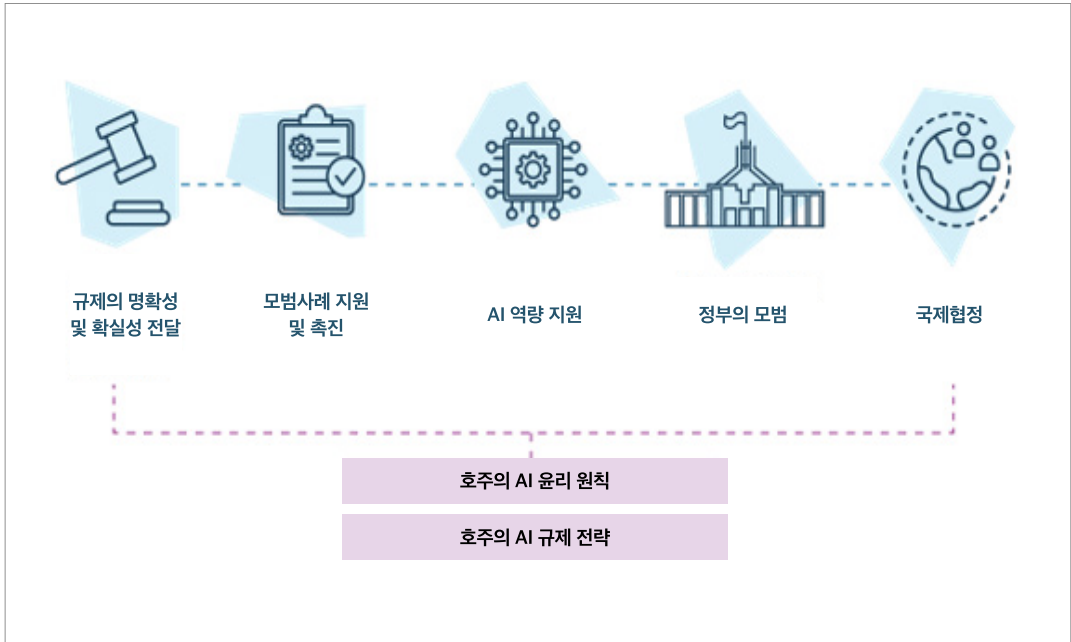
7) United Kingdom, Oral Statement to Parliament, The King's Speech 2024, 17 July 2024. <https://www.gov.uk/government/speeches/the-kings-speech-2024> Accessed 12 November 2024.

8) US National Conference of State Legislatures, Summary: Artificial Intelligence 2024 Legislation, USA Government, 3 June 2024.

9) Parliament of Canada, Minister of Innovation, Science and Industry, Amendments to Bill C-27 (44-1), 28 November 2023.

10) United Nations, General Assembly Adopts Landmark Resolution on Artificial Intelligence, UN News, 21 March 2024. <https://news.un.org/en/story/2024/03/1147831#:~:text=21%20March%202024%20SDGs%20The%20UN%20General%20Assembly,that%20will%20also%20benefit%20sustainable%20development%20for%20all.>

<그림 1> 호주 AI 법률의 목표

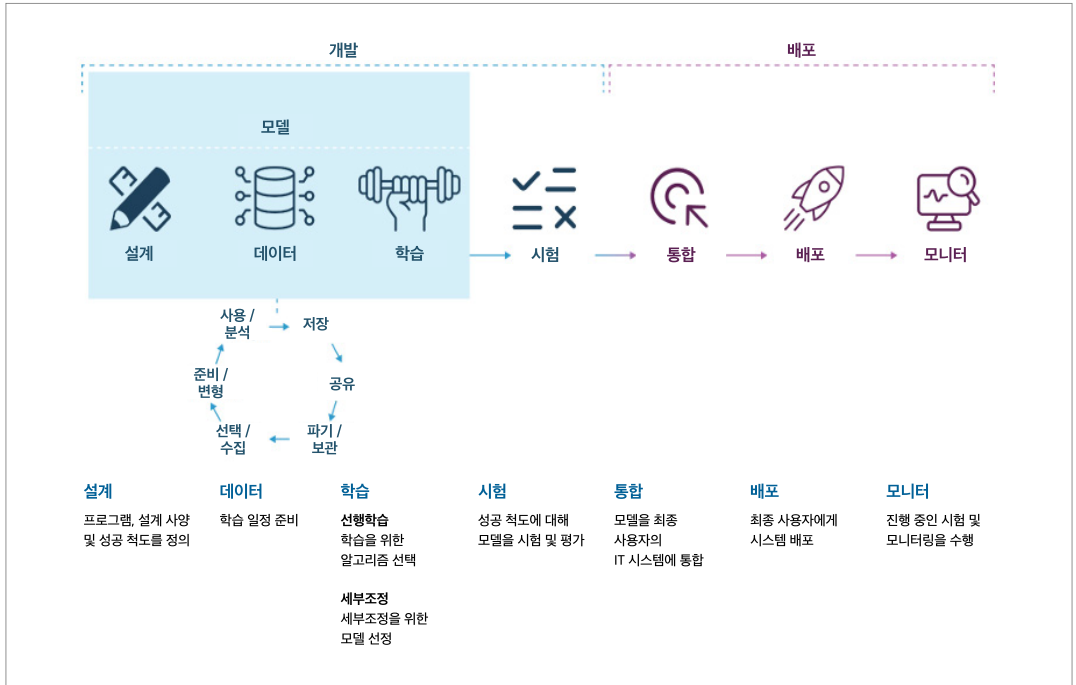


출처: 호주 산업과학자원부, 호주의 안전하고 책임 있는 AI: 고위험 환경에서 AI에 대한 의무적 보호장치를 도입하기 위한 제안, 2024년 9월, p.6

2.2 시행 범위 제안

- 2024년 고위험 환경에서 AI에 대한 의무적 보호장치를 도입하기 위한 제안서에서 제안한 보호장치는 AI설계, 시험 및 배포의 전체 생애주기에 걸쳐 적용된다. 이러한 AI 생애주기는 아래의 그림 2와 같다.

<그림 2> AI 생애주기

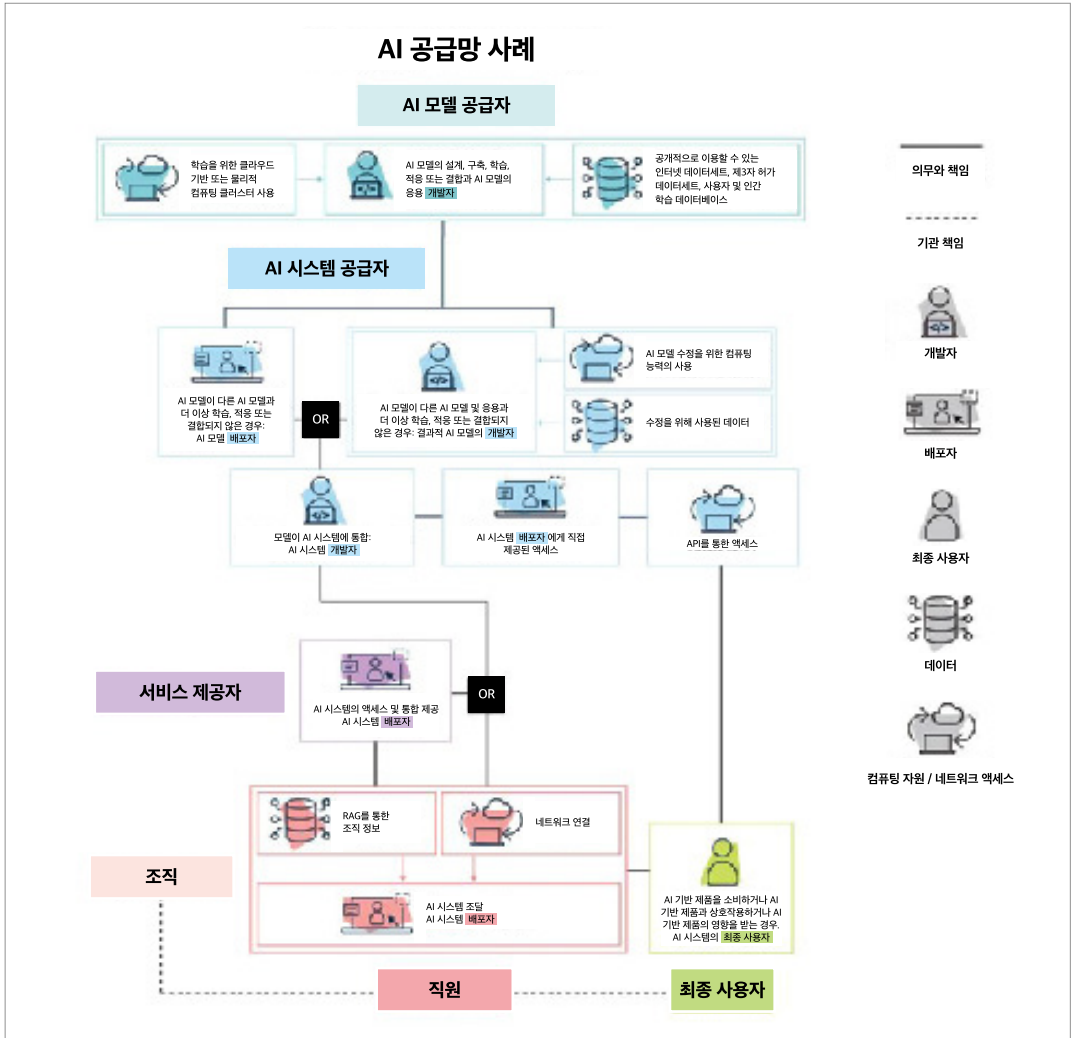


출처: 호주 산업과학자원부, 호주의 안전하고 책임 있는 AI: 고위험 환경에서 AI에 대한 의무적 보호장치를 도입하기 위한 제안, 2024년 9월, p.10

- 또한 의무적 보호장치는 AI 모델 공급자부터 AI 시스템 공급자, 서비스 제공자, 조직, 직원, 최종 사용자에게 이르기까지 AI 공급망 전반에 적용된다. 제안된 법률이 적용되는 예는 아래의 그림 3에서 설명한다.

2.3 AI 공급망의 규제

<그림 3> AI 공급망



출처: 호주 산업과학자원부, 호주의 안전하고 책임 있는 AI: 고위험 환경에서 AI에 대한 의무적 보호장치를 도입하기 위한 제안, 2024년 9월, p.34

2.4 호주 AI 전략

- 마지막으로, 이 문서는 제안된 포괄적 AI 규제 전략의 주요 요소를 요약한다. 아래의 그림 4에서는 호주 정부의 AI 규제 전략 중 주요 요소를 기술한다.

<그림 4> 호주의 AI 규제 접근법

정부 AI 규제 전략의 주요 요소	
명확한 규제 목표 및 원칙 설정	<p>예:</p> <ul style="list-style-type: none"> • 위해로부터 국민의 보호 • 경제적 편익을 위한 혁신 • 사회 복지
규제 접근법 명확화	<p>예:</p> <ul style="list-style-type: none"> • 위험 기반 접근법의 적용 • 위해의 정의, AI 위험의 종류 및 수준 명시 • 인권의 근거 구축
호주의 고유한 환경에 부합하는 규제 접근법 확보	<p>다음 사항을 고려한다.</p> <ul style="list-style-type: none"> • 경제 번영, 국가 안보 및 사회 통합을 포함한 호주의 국익에 영향을 미치는 확인된 핵심 기술로서의 AI • 호주의 경제적 이익 및 비교 우위
정부의 규제 접근법 상 일관성 확보	<p>다음 사항을 포함한 공공 및 민간 부문 접근법의 일치성 확보 및 의무 명확화</p> <ul style="list-style-type: none"> • 의무적 보호장치의 설치 • 현행 법률의 강화 및 명확화 • 자발적 AI 안전 표준 • 정부의 모범
법률 용어 및 규칙의 명확성과 일관성 촉진	<p>예:</p> <ul style="list-style-type: none"> • AI 개발, 배포 및 사용에 적용되는 현행 법률의 법률용어의 일관성과 부문 간 명확성 개선 • 법적 책임의 명확화 • 주 및 준주 법률과의 상호작용 명확화

정부 AI 규제 전략의 주요 요소	
현행 법률 및 신법의 적용 명확화	<ul style="list-style-type: none"> • 현행 법률(프라이버시, 소비자보호, 지적재산 등)의 명확화 및 강화 • 현행 법률과 새로운 의무적 보호장치 간의 상호작용 명확화 • 부문별 규제기관의 역할 명확화
경성법 및 연성법 조치의 목적 명확화	<p>다음 법률의 역할을 고려한다.</p> <ul style="list-style-type: none"> • AI 시스템 개발자 및 배포자의 행위에 대한 최소기대치를 정하는 경성법 • 추가 지침을 제공하고 산업 지배구조 관행의 이행을 지원하는 연성법
규제 이니셔티브의 이해관계자 협의의 접근법 명시	<p>다음 사항을 포함한다.</p> <ul style="list-style-type: none"> • 산업, 조합, 시민사회 및 학계와의 진실된 협의 진행 • 잠재적으로 영향을 받는 공동체 및 시민사회의 대표자가 참여하는 포괄적 접근법 • 법률, 정책 및 기술 지식을 포함한 독립적 전문가 자문
효과적이고 즉각적인 규제 생태계를 촉진하기 위해 필요한 환경의 결정	<p>다음 사항을 고려한다.</p> <ul style="list-style-type: none"> • 법률을 집행하기 위해 필요한 규제기관의 역량, 자원 및 권한 • 법적 및 윤리적 책임을 이해하고 유지하기 위한 실무지침 • 개인 권리의 행사를 지원하는 교육 및 자원 • 중소기업의 의무 이해 및 충족 지원
국제적 조정 및 상호운용성 고려	<p>호주의 규제 접근법이 다음 사항을 고려하도록 한다.</p> <ul style="list-style-type: none"> • 특히 국경을 넘어 운영하는 호주 기업의 국제적 조정 및 상호운용성 • 인권에 대한 AI 법률의 국제 규범 근거

출처: 호주 산업과학자원부, 호주의 안전하고 책임 있는 AI: 고위험 환경에서 AI에 대한 의무적 보호장치를 도입하기 위한 제안, 2024년 9월, p.59 부속서 C

3. 불확실성 지속 문제

- 제안된 고위험 환경에서의 AI 의무적 보호장치에는 여러 가지 불확실성이 있다.
- 첫째, ‘고위험 환경’을 구성하는 것이 무엇인지에 관해 몇 가지 불확실성이 있다. 이와 관련하여 이 문서는 두 가지 잠재적인 고위험 AI 범주를 기술한다. 첫 번째 범주는 특정한 AI 시스템이나 모델을 사용한다는 사실을 알고 있거나 예측할 수 있는 상황과 관계가 있다. 이 경우, 사실적 맥락과 관련하여 위험을 평가할 수 있다. 두 번째 범주는 합리적으로 잠재적 위험을 예측할 수 없는 AI 시스템 및 모델과 관계가 있다. 이 문서는 특정한 AI 시스템이 고위험인지 여부를 판단하기 위하여 사용할 수 있는 여러 원칙을 기술한다. 이러한 위험에는 ‘정당한 사유 없이 호주의 인권법에서 인정한 개인의 권리에 부정적 영향을 미칠 위험¹¹⁾’, ‘개인의 신체 또는 정신 건강이나 안전에 부정적 영향을 미칠 위험¹²⁾’, ‘개인에게 부정적인 법적 영향, 명예훼손 또는 이와 유사한 중요한 영향을 미칠 위험¹³⁾’, ‘개인의 집단에 부정적 영향을 미치거나 문화적 집단의 집단적 권리에 부정적 영향을 미칠 위험¹⁴⁾’, ‘광범위한 호주의 경제, 사회, 환경 및 법치주의에 부정적 영향을 미칠 위험¹⁵⁾’ 등이 있다. 다만, 제안서는 이러한 위험의 식별 및 평가에 관한 위험평가체계를 제공하지는 않는다. 더구나 ‘광범위한 호주의 경제, 사회, 환경 및 법치주의’에 부정적 영향을 미칠 위험 등의 원칙은 매우 광범위하고, 적용이 불확실하여 잠재적 불일치를 초래하게 될 것이다.
- 둘째, 제안된 의무적 보호장치를 현행 법률과 일치시킬지 명확하지 않다. 호주는 AI 특정 법률을 광범위하게 시행하고 있지 않지만, 현재 다수의 일반 법률이 적용되고 있다. 예를 들어, 생성형 AI 자문에 대해서는 부주의한 허위진술에 대해 확립된 법률이 적용되는 것과 같이 거짓 또는 부정 표시 금지 등의 소비자 법률이

11) Australian Government, Department of Industry, Science and Resources, Safe and Responsible AI in Australia Proposals Paper for Introducing Mandatory Guardrails for AI in High-risk Settings, September 2024, p. 19, 원칙 a.

12) 위 각주, 원칙 b.

13) 위 각주, 원칙 c.

14) 위 각주, 원칙 d.

15) 위 각주, 원칙 e.

적용되고 있다. 앞으로 나아갈 방향을 계획하면서 이 문서는 개선을 위한 선택사항을 제시하고 있다. 이러한 선택사항에는 필요한 경우 현행 규제 체계 내에서 보호장치를 채택하는 것과 관련된 '특정 분야 중심 접근법', 현행 법률을 개정하기 위해 새로운 AI 특정 법률을 제정하는 것과 관련된 '기본 접근법' 또는 예를 들어 AI 법 등의 새로운 AI 특정 법을 제정하는 것과 관련된 '경제 전체 접근법' 등이 있다. 그러나 이 문서는 AI에 적용될 수 있는 현행 법률을 보여주거나 그 운영 순서를 설명하려고 하지는 않는다. 2024년 영국의 AI 규제원칙 이행 보고서에서 영국이 시행한 단계가 분석을 진행하는데 도움이 되었을 것이다.¹⁶⁾

- 마지막으로, 법률을 모니터링하고 불법행위를 기소하기 위해 어떤 집행 및 행정 기관과 기반시설이 사용될지는 명확하지 않다. AI의 투명하지 않은 특성에 비추어 그 운영과 영향을 판별하기 어려운 경우가 많다. 또한 AI의 보편적 특성을 고려할 때 AI 사용을 감시하고 불법행위를 기소하는 것은 비용이 많이 들고 논리적으로 어렵다. 이러한 맥락에서 AI 규제의 새로운 접근법이 필요하다고 제안한다. 필자는 계획적 준수 접근법을 제안한다. 이 문서의 다음 절에서 이 접근법을 중점적으로 다루기로 한다.

4. 법률에서 계획적 준수로의 이행

- 계획적 준수는 AI 알고리즘, AI 톨 및 AI 통제 내 입법 명령과 기준을 포함시키는 것과 관계가 있다. 이 개념은 '계획적 준수를 통한 합법적인 AI 발전'¹⁷⁾이라는 필자의 논문에서 먼저 제시하였다. 제안한 계획적 준수 체계는 AI 시스템을 설계, 시험 및 배포하는 기술 변호사와 컴퓨터 과학자 간의 긴밀한 학제 간 협력이 필요하다. 또한 이 체계는 AI 위험을 평가하고 특정한 AI 시스템이나 모델의 위험 특성과 수준을 조정하는 규제의 설계는 지원하는 위험평가 전문가가 관여할 필요가 있다. 또 이 체계는 AI의 잠재적 생산성과 효율성 증대와 위험을 적절히 조정하는 것과 관계가 있다. 마지막으로, 특정한 AI 시스템이 산업 또는 사회적 집단 내에서 어떻게 작용할지에 대해 조언할 수 있는 공공부문, 건강 및 재정 전문가 등 분야별 전문가의 관점을 통합하는 것이 매우

16) United Kingdom, Department for Science Innovation & Technology, Implementing the UK's AI Regulatory Principles Report (2024).

17) Niloufer Selvadurai, 'Advancing AI through "Compliance by Design"' (2025) 31(1) Computer and Telecommunications Law Review 1; 이 개념은 블록체인 스마트계약 규제의 맥락에서 Niloufer Selvadurai, 'Mitigating the legal challenges associated with blockchain smart contracts: the potential of hybrid on-chain/off-chain contracts,' (2023) 80(3) Washington & Lee Law Journal 1163-1179에서 광범위하게 적용되고 있다.

중요하다. 계획적 준수 체계는 법률을 대체하는 것이 아니라 법률의 중요한 목적을 달성하고 법률을 효과적으로 집행하는데 도움이 되는 보충적 체계라는 점을 주의해야 한다.

결론

- 그러므로 전 세계의 법률 및 정책 입안자가 AI의 생산성 및 경제적 편익을 지원하면서 그 합법적이고 책임 있는 사용을 지원하기 위해 AI를 규제하는 최선의 방법을 고려하고 있는 것처럼, 호주의 고위험 환경에서 AI에 대한 의무적 보호장치를 도입하기 위한 제안은 미래의 법률을 계획할 때 다룰 필요가 있는 문제에 대한 유용한 시각을 제공한다. AI 시스템과 모델의 설계 내에 준수를 포함시키고 법률의 효과적 집행을 지원하기 위해 ‘계획적 준수’ 체계가 AI 법률을 뒷받침해야 한다.

AI 법률에서 컴플라이언스 설계로의 전환: 호주의 관점

발행일 2024년 11월 30일

발행인 한영수

발행처 한국법제연구원

세종특별자치시 국책연구원로 15(반곡동, 한국법제연구원)

T. 044)861-0300 F.044)861-9915

등록번호 1981. 8. 11. 제2014-000009호

<http://www.klri.re.kr>

- 본원의 승인없이 轉載 또는 譯載를 禁함.
- 이 책자의 내용은 본원의 공식적인 견해가 아님.

본원의 공식 EU AI ACT 번역 내용은 [AI혁신법제연구사업 번역자료집]

EU AI ACT 번역본, (2024. 7. 30.) 을 참고해주시기 바랍니다.

(한국법제연구원 - 자료집 - 806)

Artificial Intelligence Law Issue Brief