

# 국 가 인 권 위 원 회

## 결 정

제 목 인공지능 개발과 활용에 관한 인권 가이드라인

### 주 문

국가인권위원회는 「국가인권위원회법」 제25조 제1항의 규정에 따라 다음과 같이 권고합니다.

#### 1. 국무총리에게

[붙임]의 「인공지능 개발과 활용에 관한 인권 가이드라인」에 기초하여 인공지능 관련 정책이 수립·이행되고, 관계 법령이 제·개정되도록 관련 부처들을 유기적으로 조정하고 통할할 것을 권고합니다.

#### 2. 과학기술정보통신부장관, 개인정보보호위원회 위원장, 방송통신위원회 위원장, 공정거래위원회 위원장, 금융위원회 위원장에게

가. [붙임]의 「인공지능 개발과 활용에 관한 인권 가이드라인」에 기초하여 인공지능 관련 정책을 수립·이행하고, 관계 법령을 제·개정할 것을 권고합니다.

나. 공공기관 및 민간기업이 [붙임]의 「인공지능 개발과 활용에 관한 인권 가이드라인」을 준수하여 인공지능으로 인한 인권침해 및 차별이 발생하지 않도록 적극적으로 관리·감독할 것을 권고합니다.

# 이 유

## I. 권고의 배경

인공지능(AI: Artificial Intelligence) 기술의 빠른 발전은 고용, 금융, 행정, 복지 등 사회 전반에 걸쳐 인간의 기본적인 삶과 인권에 영향을 미치고 있다.

인공지능의 발전과 확산으로 인해 개인 삶의 질이 높아질 것으로 기대되지만, 안면인식기술이나 인공지능에 의한 채용, 플랫폼 노동, '이루다' 사태에서 보듯이 개인정보 및 사생활에 관한 권리 침해, 차별 등과 같은 인권 문제도 발생하고 있다.

그러나 인공지능으로 영향을 받는 당사자들은 인공지능의 도입, 운영, 결정에 대하여 의견 제시와 참여의 기회를 보장받고 있지 못하며, 인공지능에 의해 인권 침해가 발생한 경우에도 효과적인 권리구제를 받을 수 있는 절차와 방법이 미흡한 상황이다.

따라서 인공지능을 개발할 때부터 개인의 삶과 사회적 공익에 기여할 수 있도록 개발하고, 인간의 존엄성과 자기결정권의 보장 및 차별받지 않을 권리 등 인권에 기반하여 활용하며, 인공지능 기술로 발생할 수 있는 인권침해의 예방 및 권리구제 절차를 마련하는 것이 중요하다.

이에, 국가인권위원회는 「국가인권위원회법」 제25조 제1항에 따라 인권적 관점에서 인공지능의 개발과 활용 시 준수해야 할 내용을 다음과 같이 검토하였다.

## II. 판단 및 참고기준

- 「대한민국헌법」 제10조, 제11조, 제17조, 제18조
- 「경제적·사회적 및 문화적 권리에 관한 국제규약」 제15조
- EU의 인공지능 가이드라인(2019), OECD의 인공지능 권고(2019) 등

### Ⅲ. 판단

#### 1. 인공지능에 관한 국제 인권 기준

인공지능에 관한 여러 국제 인권 기준은 ‘인권 및 인간 존엄성의 존중’을 최우선적 가치로 명시하고 있고, 인공지능을 개발하고 활용하는 모든 이해관계자는 인공지능의 개발, 배치, 활용 등 전 과정에서 인권을 존중하고 보호해야 하며, 이에 위배되는 결과에 대해 적절한 보호수단을 강구하도록 하고 있다.

또한 인공지능이 부정적인 편견이나 고정관념을 학습한 결과로 차별을 조장하거나 지속시킬 위험성이 있음을 지적하고, 인공지능이 사용하는 데이터의 편향, 불완전성 등에 대한 사전적 식별과 조치를 강조하고 있다.

국제 인권 기준은 인공지능 시스템에 대한 ‘설명가능성과 책임성의 확보’를 중요시하고 있다. 설명가능성 또는 투명성이란 인공지능 시스템에 의해 내려진 판단이나 결정에 대해 그 이유, 과정, 결과 등을 인간이 이해하고 설명을 들을 수 있어야 한다는 것이고, 책임성은 인공지능의 판단이나 결정에 대해 책임을 질 수 있는 주체와 절차를 명확히 하고 그에 맞는 법적·제도적 근거를 갖출 것을 요구하는 것이다.

일부 인공지능 기반의 시스템은 대량의 개인정보 수집 및 활용 기술과 관련이 있고, 개인의 민감한 정보나 사생활을 추적·감시하는 데 활용될 여지가 있어 ‘개인정보 및 사생활의 보호’를 강조하고 있으며, 동시에 인공지능 알고리즘의 보안과 안전성 확보가 중요하다.

#### 2. 인권 가이드라인의 필요성

인공지능으로 인해 발생할 수 있는 문제를 예방하기 위하여 국제사회는 인공지능의 개발과 활용에 있어 인권에 대한 보호 기준이 마련되어야 한다는 데 인식을

같이하고, 각 분야에서 인공지능 관련 제도와 지침을 개발하고 있다.

반면 우리나라의 법률 및 제도는 선진국과 비교해 인공지능 기술 및 산업의 성장에 중점을 두어 인공지능으로 인한 인권 문제를 다루는 데 부족한 측면이 있고, 각 정부 부처와 기관은 각자의 소관 업무와 관련하여 개별적 이슈로 접근하고 있다.

인공지능과 관련하여 총체적인 인권적 접근이 이루어지지 못하면 표현의 자유, 혐오와 차별, 개인정보 보호 등 사회적 관심을 받는 주제는 더욱 주목을 받게 되지만, 정보 및 서비스 접근성, 개인의 선택권 및 자기결정권 등 중심 주제에서 벗어나 있는 문제는 해결이 더욱 어려워질 수 있다.

또한 각 정부 부처에서 발표하는 다양한 인공지능 관련 윤리 기준은 역사적 배경과 문화적 차이에 따라 자의적으로 해석될 수 있어, 국제적으로 인정되고 국가의 법률 등에서 명확히 밝히고 있는 인권 기준과는 의미와 규범성이 다르다고 할 수 있다.

인공지능의 개발로 인해 예상치 못한 사회적 문제점이 점차 증가할 것을 고려할 때 인공지능의 개발과 활용에 있어 인권적 가치를 훼손하지 않고, 인간의 존엄성과 기본권 보장을 위한 인권 가이드라인의 필요성이 커지고 있다.

따라서 인권 가이드라인은 기존의 제도나 기준이 가지는 한계를 극복하고, 인공지능이 영향을 미치는 영역에 대해 총체적인 접근과 명확한 기준을 제시할 수 있다는 점에서 그 필요성이 더 크다고 할 수 있다.

### 3. 인권 가이드라인의 주요 내용

#### 1) 인간의 존엄성

「대한민국헌법」 제10조에서 보장하고 있는 인간의 존엄과 가치는 누구나 누려야 할 불가침의 기본적 인권으로, 어떠한 활동도 인간의 존엄에서 유래하는 다양한 권리의 희생을 강요해서는 안 되며, 모든 활동은 인간의 존엄과 가치를 향상시키는 방향으로 수행되어야 한다.

인공지능은 인간의 존엄과 가치 및 행복을 추구할 권리에 부합하는 방향으로 개발 및 활용되어야 하며, 개인의 선택과 판단 및 결정을 강요하거나 자율성을 침해할 수 없다.

## 2) 투명성과 설명 의무

인공지능 기술을 활용한 판단 과정과 그 결과에 대해 적절하고 합리적인 설명이 보장되어야 한다. 공공기관이 인공지능을 통한 의사결정을 할 때 설명 가능한 인공지능이 활용되어야 하며, 특히 조달의 경우 입찰 단계부터 설명가능성이 보장되어야 한다.

공공기관이 개발하고 활용하는 모든 인공지능과 민간이 개발하고 활용하는 인공지능 중 개인의 생명이나 안전 등 기본적 인권에 중대한 영향을 미치는 인공지능은 사용된 데이터와 인공지능 알고리즘의 주요 요소를 일반에게 공개하고 설명하여야 한다.

특히 완전히 자동화된 의사결정으로만 개인에게 법적 효력 또는 생명, 신체, 정신, 재산에 중대한 영향을 미치는 일은 제한되어야 하고, 이러한 의사결정이 이루어진 경우에는 당사자가 해당 방식을 거부하거나 인적 개입을 요구할 수 있는 권리를 보장해야 한다.

## 3) 자기결정권의 보장

개인정보자기결정권은 정보주체의 자기 정보에 대한 통제력을 보장하기 위한 것이고, 가장 중요한 정보주체의 동의는 단순한 외형적 의사 표시만이 아니라 정보주체가 개인정보 처리에 대한 제반 상황을 충분히 설명받고 자유의사에 따라 결정하는 것을 의미한다.

정보주체는 인공지능 서비스가 언제, 어디서 자신의 개인정보를 수집하고, 어떻게 개인정보를 처리하여 사용, 보관, 삭제하는지에 대해 알고 참여할 권리가 있다.

인공지능의 개발과 활용에서 개인정보는 목적에 필요한 최소한의 범위에서 처리되고, 처리목적 달성에 필요한 기간 동안만 보관되어야 한다.

인공지능을 활용한 의사결정의 내용과 관련성이 없거나, 부정확한 데이터에 기반한 의사결정이 이루어지지 않도록 데이터의 정확성, 완전성, 최신성을 보장해야 한다.

#### 4) 차별 금지

인공지능을 개발하고 활용할 때는 인공지능으로 인해 영향을 받는 사람의 다양성과 대표성을 반영하기 위해 노력해야 하고, 성별, 종교, 장애, 나이, 출신 지역, 신체조건, 피부색, 성적 지향, 사회적 신분 등 개인과 집단의 특성에 따라 편향적이고 차별적인 결과가 나오지 않아야 한다.

데이터의 수집 및 시스템 설계, 활용 등 인공지능 개발 전반에 걸쳐 편향이나 차별을 배제해야 하고, 데이터 요소를 검사하고 차별적인 데이터를 조정하는 등의 조치를 해야 한다.

개발한 인공지능에 대해 주기적인 모니터링을 거쳐 데이터 품질과 위험을 관리하고, 차별적 결과나 의도하지 않은 결과에 대해 개선의 조치를 주기적으로 수행하여야 한다.

#### 5) 인공지능 인권영향평가 시행

국가는 인공지능의 개발과 활용에 있어서 인권침해와 차별의 가능성 및 정도, 영향을 받은 당사자의 수, 사용된 데이터의 양 등을 고려하여 인권영향평가를 실시해야 한다.

인공지능 인권영향평가 내용에는 인공지능의 특성, 상황, 범위 및 목적을 감안하여 인권 가이드라인이 제시한 원칙 및 내용, 국제 인권 기준, 관련 법률에서 정한 의무 등이 포함되어야 하며, 인권침해 위험요인의 분석 및 개선 사항 등을

도출해야 한다.

인공지능 인권영향평가 결과에서 인권에 미치는 부정적인 영향이나 편향성 및 위험성이 드러난 경우 이를 방지하거나 완화하기 위한 조치사항을 수립하여 적용하여야 하며, 원칙적으로 그 내용이 공개되어야 한다.

## 6) 위험도 등급 및 관련 법·제도 마련

국가는 공공기관과 민간이 개발하고 활용하는 인공지능이 인권과 책임성을 보장할 수 있도록 관련 법률과 감독 체계를 보완해야 하며, 구체적인 지침과 정책을 마련해야 한다.

국가는 개인의 인권과 안전에 미치는 위험성이 매우 높아 인공지능이 금지되는 영역, 상당한 제한이 필요한 인공지능 고위험 영역, 일정 정도의 제한이 필요한 제한적 위험성을 가진 영역, 위험성이 거의 없는 영역 등 적절하게 인공지능의 위험성 단계를 구분하고, 그에 맞는 규제 수준과 인적 개입이 이루어지도록 법과 제도를 마련하여야 한다.

국가는 인공지능을 독립적이고 효과적으로 감독할 수 있는 체계를 수립하여 개인의 인권과 안전을 보장하고, 인공지능에 의해 피해를 입은 사람이 진정을 제기하는 등의 방법으로 권리 구제를 받을 수 있는 기회를 제공하여야 한다.

## IV. 결론

위와 같은 이유로 「국가인권위원회법」 제25조 제1항의 규정에 따라 주문과 같이 권고하기로 결정한다.

[붙임] : 「인공지능 개발과 활용에 관한 인권 가이드라인」

2022. 4. 11.

위 원 장 송 두 환

위 원 이 상 철 <불참>

위 원 박 찬 운

위 원 남 규 선

위 원 문 순 회

위 원 이 준 일

위 원 석 원 정

위 원 서 미 화

위 원 윤 석 희

위 원 김 수 정

위 원 한 석 훈



[붙임] 인공지능 개발과 활용에 관한 인권 가이드라인

## 인공지능 개발과 활용에 관한 인권 가이드라인

### 제1장 개요

#### 제1절 제정 배경

1. 사회 전반에 걸쳐 인공지능(AI: Artificial Intelligence)의 활용이 증가함에 따라 고용, 금융, 행정, 복지 등 거의 모든 분야에서 인간의 기본적인 삶과 인권에 영향을 미치는 사례가 증가하고 있습니다.
2. 앞으로도 막대한 양의 빅데이터를 바탕으로 다양한 영역에서 사람의 판단을 대신할 수 있는 인공지능은 점차 적용영역을 넓혀 사회 전반과 개인의 삶에 강력한 영향력과 파급력을 행사할 것입니다.
3. 인공지능의 발전과 확산은 국가경쟁력과 개인 삶의 질을 높일 것으로 기대되지만 개인정보 및 사생활에 대한 권리, 차별받지 않을 권리 등 인권을 침해하는 문제들도 대두되고 있습니다.
4. 반면 인공지능으로 영향을 받는 당사자들은 인공지능의 도입, 운영, 결정에 대하여 참여의 기회를 보장받고 있지 못하며, 인공지능으로 인한 인권 침해가 발생한 경우에도 적절하고 효과적인 권리구제를 받을 수 있는 절차와 방법이 미흡한 상황입니다.
5. 따라서 인공지능을 개인의 삶과 사회적 공익에 기여할 수 있도록 설계 하며, 인간의 존엄성과 자기결정권 및 차별을 받지 않을 권리의 보장 등 기본적인 인권에 기반을 두도록 하는 것이 매우 중요합니다.
6. 본 가이드라인은 인권적 관점에서 인공지능의 개발과 활용의 전 과정에서 인권침해를 예방하기 위하여 준수해야 할 원칙과 내용을 제시하고자

마련되었습니다.

7. 국가인권위원회는 「국가인권위원회법」 제19조 제6호에 따라 인권침해의 유형, 판단기준 및 그 예방조치 등에 관한 지침을 제시할 권한을 가지고 있으므로, 인공지능에 의한 인권침해와 차별을 판단하고, 개선을 권고하거나 구제절차를 마련하는 데 필요한 기준을 제시하고자 합니다.

## 제2절 목적과 의미

8. 본 가이드라인은 인공지능의 개발과 활용에 있어서 우리 사회의 인권적 가치가 훼손되지 않고 인간의 존엄성을 보장하며 기본적 인권을 실현하는 방향으로 나아가도록 하는 데 목적이 있습니다.
9. 또한 인공지능의 개발과 활용 과정에 적용할 인권원칙을 제시하고, 인공지능 서비스 이용자, 영향을 받는 당사자 등에게 주어진 권리 및 피해 구제수단을 제공하며, 정부에게 적절한 법령과 정책을 수립할 수 있는 가이드라인을 마련하고자 합니다.
10. 본 가이드라인은 인공지능의 개발부터 활용에 이르는 모든 과정에 참여하는 사회구성원들을 대상으로 하며, 여기에는 인공지능 개발자, 이용자, 영향을 받는 당사자, 정부 및 공공기관, 기업 등이 포함됩니다.

## 제3절 정의

11. 인공지능의 개발과 활용을 위한 가이드라인에서 의미하는 인권은 「대한민국헌법」과 법률에서 보장하거나 대한민국이 가입·비준한 국제인권조약 및 국제관습법에서 인정하는 인간으로서의 존엄과 가치 및 자유와 권리를 의미합니다.
12. 본 가이드라인에서 말하는 인공지능은 일차적으로는 학습과 추론, 판단을 전자적으로 구현하는 알고리즘과 해당 프로세스를 지칭하나, 이차적으로는 빅데이터 등 인공지능을 기능하게 하는 일련의 기술들을 포함

하고 있습니다.

13. ‘영향을 받는 당사자(affected individuals)’ 는 국가나 기업의 규정 또는 행위 등으로 인하여 인공지능의 적용대상이 되고, 직·간접적으로 인권에 영향을 받는 개인 또는 집단을 의미합니다.

## 제2장 가이드라인 내용

### 제1절 인간의 존엄성 존중

14. 「대한민국헌법」 제10조에서 보장하고 있는 인간의 존엄과 가치는 누구나 누려야 할 불가침의 기본적 인권으로, 모든 권리의 출발점인 동시에 종국적으로 보장되어야 할 인권적 가치입니다.
15. 어떠한 활동도 인간의 존엄에서 유래하는 다양한 권리들의 희생을 강요해서는 안 되며, 궁극적으로 모든 활동은 인간의 존엄과 가치를 향상시키는 방향으로 수행되어야 합니다.
16. 인공지능은 인간으로서의 존엄과 가치 및 행복을 추구할 권리에 부합하는 방향으로 개발 및 활용되어야 하며, 개인의 선택과 판단 및 결정을 강요하거나 자율성을 침해해서는 안 됩니다.
17. 인공지능의 개발과 활용은 개인의 행복과 사회적 공공성의 증진에 위배되어서는 안 되고, 표현의 자유, 집회 및 결사의 자유, 노동권 등 인권을 인공지능의 부정적 영향으로부터 보호해야 합니다.

### 제2절 투명성과 설명 의무

18. 알 권리란 개인이 의사형성에 필요한 정보를 수집하고, 수집된 정보를 취사·선택할 수 있는 자유를 의미하며, 알 권리를 통해 개인은 충분한 정보를 획득할 수 있고, 지식과 이해의 폭을 넓혀 합리적인 판단과

자신의 인격을 발현시킬 수 있습니다.

19. 이러한 알 권리의 보장과 인공지능이 미치는 영향력과 중요성을 감안할 때, 인공지능의 판단과정과 그 결과에 대한 적절하고 합리적인 설명이 보장되어야 합니다. 학습 및 추론, 판단의 과정과 결과에 이른 이유를 설명하기 어려운 인공지능은 이에 대한 대응의 불확실성과 영향을 받는 당사자의 불안감을 유발하고, 인권 및 안전에 관한 법령과 정책의 집행 효과를 불분명하게 할 수 있습니다.
20. 인공지능과 상호작용하는 사람에게는 언제나 그 상대방이 인공지능이라는 사실을 알려야 합니다.
21. 공공기관은 인공지능의 개발과 활용 계획 등을 사전 공개하여야 하고, 관련 당사자들의 의견을 공청회 등으로 수렴하여야 합니다. 공공기관이 인공지능을 통한 의사결정을 할 때 설명할 수 없는 인공지능을 활용해서는 안 되며, 특히 조달의 경우 입찰 단계에서부터 설명가능성이 보장되어야 합니다.
22. 공공기관이 개발하고 활용하는 모든 인공지능과 민간이 개발하고 활용하는 인공지능 중 개인의 생명이나 안전 등 기본적 인권에 중대한 영향을 미치는 인공지능은 사용된 데이터와 인공지능 알고리즘의 주요 요소 일반에게 공개하고 설명하여야 합니다.
23. 또한, 인공지능에 의한 자동화된 의사결정이 예정되어 있는 경우, 영향을 받는 당사자들은 사전에 그 사실을 알아야 합니다. 자동화된 의사결정에 의하여 영향을 받는 당사자는 그 결정의 이유에 대하여 설명을 듣고, 당사자 진술을 할 수 있으며, 이의를 제기할 수 있어야 합니다.
24. 특히, 완전히 자동화된 의사결정으로만 개인에게 법적 효력 또는 생명·신체·정신·재산에 중대한 영향을 미치는 일은 제한되어야 하고, 이러한 의사결정이 이루어진 경우에는 당사자가 해당 방식을 거부하거나 인적 개입을 요구할 수 있는 권리를 보장받아야 합니다.

### 제3절 자기결정권의 보장

25. 「대한민국헌법」 제10조가 규정하고 있는 인간의 존엄과 행복추구권에서 도출되는 자기결정권은 결정의 주체인 개인이 중대한 사안에 대해 외부의 강요 없이 누구나 스스로의 판단에 따라 타인의 간섭 없이 결정하고 행동할 수 있으며, 동시에 선택하지 않을 자유를 의미합니다.
26. 개인정보자기결정권은 정보주체의 자기 정보에 대한 통제력을 보장하기 위해 인정되는 것인데, 그 통제력 보장의 핵심은 정보주체의 동의권입니다. 따라서 개인정보 처리에 대한 정보주체의 동의는 단순한 외형적 의사 표시만이 아니라 정보주체가 개인정보 처리에 대한 제반 상황을 설명받고 자유의사에 따라서 결정될 수 있어야 합니다.
27. 인공지능과 관련하여 정보주체의 권리는 처리된 개인정보에 관하여 고지를 받을 권리, 개인정보 접근 및 열람권, 개인정보처리 동의권 및 정정·삭제권, 처리정지요구권 등을 포함하며, 정보주체는 자신의 개인정보가 사용되는 방법을 이해하고 그에 대한 통제권을 가지는 것이 중요합니다. 정보주체는 인공지능 서비스가 언제, 어디서 자신의 개인정보를 수집하고, 어떻게 개인정보를 처리하여 사용, 보관, 삭제하는지에 대해 알고 참여할 권리가 있습니다.
28. 인공지능의 개발과 활용에서 개인정보는 목적에 필요한 최소한의 범위 내에서 처리되어야 하며, 처리목적 달성에 필요한 기간 동안만 보관되어야 합니다. 또한 이러한 개인정보 처리 방침은 정보주체가 확인할 수 있도록 공개되어야 합니다.
29. 인공지능의 개발과 활용에서 민감정보를 처리할 때에는 특별한 주의 기울여 보호하여야 합니다. 더불어 의사결정의 내용과 관련성이 없거나, 부정확한 데이터에 기반한 의사결정이 이루어지지 않도록 데이터의 정확성, 완전성, 최신성을 보장해야 합니다.

#### 제4절 차별 금지

30. 「대한민국헌법」 제11조는 모든 국민에게 평등권을 보장하고 있으며, 「국가인권위원회법」 제2조는 합리적인 이유 없이 성별, 종교, 장애, 나이, 출신 지역, 신체조건, 피부색, 성적 지향, 사회적 신분 등 개인 특성에 따라 특정한 사람을 우대·배제·구별하거나 불리하게 대우하는 행위를 평등권 침해의 차별행위로 정의하고 있습니다.
31. 따라서 모든 사람은 평등하고, 어떠한 이유로도 생활의 모든 영역에서 구별, 배척, 제한을 받지 아니하며, 인권의 인정과 향유를 보장받을 권리가 있습니다.
32. 인공지능을 개발하고 활용할 때는 인공지능으로 인해 영향을 받는 사람의 다양성과 대표성을 반영하기 위해 노력해야 하고, 성별, 종교, 장애, 나이, 출신 지역, 신체조건, 피부색, 성적 지향, 사회적 신분 등 개인과 집단의 특성에 따라 편향적이고 차별적인 결과가 나오지 않도록 하여야 합니다.
33. 또한 인공지능의 결정이 특정 집단이나 일부 계층에게 차별적이거나 부정적 영향을 초래하지 않기 위해 개발 단계부터 다양한 계층의 의견을 수렴하고, 차별적 결과가 발생하지 않도록 필요한 조치를 취해야 합니다.
34. 학습 데이터의 수집과 선정, 알고리즘의 설계와 활용방향 설정 등 인공지능 개발 전 과정에 걸쳐 편향이나 차별의 요소가 배제될 수 있도록 점검하는 절차를 마련해야 합니다. 여기에는 학습 데이터의 개별 요소를 검사하고 차별적 영향을 초래할 수 있는 데이터를 조정하는 등의 조치가 포함되어야 합니다.
35. 특히 학습용 데이터가 인공지능의 판단에 직접적인 영향을 미치는 상황을 고려할 때, 학습용 데이터의 수집 단계부터 차별적 요소를 통제하고 데이터 편향성을 최소화하여 인공지능을 통한 의사결정이 특정 집단에 부정적 영향을 미치지 않도록 해야 합니다.

36. 개발한 인공지능에 대해 주기적인 모니터링을 거쳐 데이터 품질과 위험을 관리하고, 차별적 결과나 의도치 않은 결과에 대해 개선의 조치를 주기적으로 수행해야 합니다.
37. 인공지능 기술 및 서비스에 대한 접근성과 인공지능이 주는 혜택은 사회적 약자를 포함하여 모든 사회구성원에게 평등하게 제공되어야 합니다. 또한 모든 사람의 인공지능에 대한 지식과 이해를 촉진시키고, 인공지능의 활용이 어려운 계층에 대한 교육 및 지원이 이루어져야 합니다.

### 제5절 인공지능 인권영향평가 시행

38. 국가는 인공지능의 개발과 활용에 있어 인권적 가치가 우선시 되도록 하여야 하며, 인공지능으로 인해 발생할 수 있는 인권침해와 차별에 대하여 사전적 또는 사후적으로 관리 감독을 할 의무가 있습니다.
39. 국가는 인공지능의 개발과 활용에 있어서 인권침해와 차별의 가능성 및 정도, 영향을 받는 당사자의 수, 사용된 데이터의 양 등을 고려하여 공공기관 및 민간기업을 대상으로 인권영향평가를 실시하여야 합니다. 특히 인공지능 기술이 적용되어 기존 제도로 관리되거나 감독될 수 없는 새로운 분야는 인권영향평가 제도를 도입해야 합니다.
40. 인권영향평가 내용에는 인공지능의 특성, 상황, 범위 및 목적을 감안하여 인권 가이드라인이 제시한 원칙 및 내용, 국제 인권 기준, 관련 법률에서 정한 의무 등이 포함되어야 하며, 인권침해 위험요인의 분석, 개선 사항 등을 도출해야 합니다.
41. 인권영향평가는 개발 및 출시 전에 실시하고 인공지능의 기능 또는 범위 변경 시 평가를 갱신하여야 합니다.
42. 인공지능을 개발 또는 활용하는 국가와 기업은 인권영향평가 결과에서 인권에 미치는 부정적인 영향이나 편향성 및 위험성이 드러난 경우 이를 방지하거나 완화하기 위한 조치사항을 수립하여 적용하여야 하며, 원칙적

으로 그 내용이 공개되어야 합니다. 또한, 이를 방지하거나 완화하는 조치를 취하기 전에는 그 개발과 활용을 중단해야 합니다.

43. 국가는 인권영향평가를 인권전문성과 독립성을 확보한 기관이 담당하도록 하고, 인권영향평가의 활성화를 위하여 관계 전문가의 육성, 영향평가 기준의 개발 및 보급 등 필요한 조치를 마련해야 합니다.

### 제6절 위험도 등급 및 관련 법제도 마련

44. 국가는 공공기관과 민간이 개발하고 활용하는 인공지능에서 인권과 책임성을 보장하도록 관련 법률과 감독 체계를 보완해야 하며, 특히 인권을 보장하기 위한 구체적인 지침과 정책을 마련하여야 합니다.
45. 국가는 개인의 인권과 안전에 미치는 위험성이 매우 높아 인공지능이 금지되는 영역, 상당한 제한이 필요한 인공지능 고위험 영역, 일정 정도의 제한이 필요한 제한된 위험을 가진 영역, 위험성이 거의 없는 영역 등 적절하게 인공지능의 위험성 단계를 구분하고, 그에 맞는 규제 수준과 인적 개입이 이루어지도록 법과 제도를 마련하여야 합니다.
46. 특히, 당사자의 개인의 생명이나 안전 등 기본적 인권에 중대한 영향을 미치는 인공지능은 투명성과 설명가능성, 개인 정보 보호, 차별 금지 등 그 규제에 있어서 공공과 민간의 구분 없이 엄격하게 적용되어야 합니다.
47. 감독 기관은 공공기관과 민간의 위법한 인공지능 개발과 활용 여부를 조사하고 피해 구제 및 조치를 취하기 위하여 상세 정보에 접근할 수 있어야 합니다. 이를 위하여 공공기관 인공지능 및 민간 고위험 인공지능 개발자 및 운영자는 사용된 데이터와 알고리즘의 주요 요소 등을 기록하고 문서화하여 일정 기간 보관하여야 합니다.
48. 국가는 인공지능을 독립적이고 효과적으로 감독할 수 있는 체계를 수립하여 개인의 인권과 안전을 보장하고 피해를 구제하여야 합니다. 이러한 인공지능 국가 감독 체계는 인지 또는 진정으로 접수한 사건을 조사



하기에 충분한 자원, 권한 및 전문지식을 구비해야 합니다.

49. 국가는 인공지능으로 인하여 인권을 침해당하거나 차별을 받은 사람이 진정을 제기하여 권리를 구제받을 수 있는 기회를 보장하는 등 국가 기관의 구제수단에 대한 접근을 제공해야 합니다. 인공지능을 개발하고 활용하는 공공기관과 민간은 언제든지 구제가 가능하도록 그 책임자에 대한 정보는 물론, 이의를 제기할 수 있는 기관과 방법에 대한 정보를 일반에 공개하여야 합니다.
50. 국가는 대량 감시와 차별로 이어지고, 집회 및 결사의 자유에 부정적 영향을 행사할 위험이 높은 얼굴인식 등 원격 생체인식 기술의 사용을 공공장소에서 원칙적으로 금지해야 합니다. 해당 기술은 특별한 경우에만 사용하되, 인권침해나 차별의 위험성이 드러난 경우 이를 방지하거나 완화하는 조치를 취하기 전까지는 사용을 중단해야 합니다.
51. 국가는 다양한 정보가 자유롭게 유통되는 정보 환경을 조성하고, 표현의 자유, 정보 접근 및 의견 표명 등에 부정적 영향을 미치는 인공지능에 대해 적절한 조치를 마련해야 합니다.
52. 국가는 확대되는 인공지능의 개발과 활용에 따라 변화가 예상되는 노동 시장의 일자리 감소, 전자적 노동 감시 등에 대한 대책 마련과 함께 이에 대한 평가와 점검을 해야 합니다.
53. 또한 국가는 생명의 존엄성 및 윤리를 훼손할 가능성이 높은 자율살상 무기에 대하여 인도주의적으로 접근하고 그 연구, 개발, 생산 및 활용을 금지하는 국제 규범을 준수하고, 이에 대한 논의에 적극적으로 참여해야 합니다.