

자율주행자동차 윤리가이드라인

2020.12

자율주행자동차 윤리가이드라인

자율주행자동차 윤리가이드라인



목 차

01

추진배경 ...1

적용범위 및
구성체계 ...7

02

03

총칙 ...11

1. 목적
2. 용어의 정의
3. 적용대상

공통원칙 ...19

1. 기본가치
2. 행동원칙

04

05

행위주체별 책임 ...31

1. 설계자
2. 제작자
3. 관리자
4. 서비스 제공자 및 이용자

결론 ...39

06

부록 ...43

01

추진배경

01 추진 배경

1. 자율주행자동차 윤리 가이드라인 제정 배경

- 자율주행자동차는 과학·기술의 급속한 발전으로 인해 인간이 직접 운전하는 전통적인 자동차에 비해 운행 시 인간의 생명과 신체의 안전을 획기적으로 개선할 것으로 예상된다.
- 그러나 자율주행 도로시험 운행 중 다양한 사고가 발생하였으며, 특히 2018년 우버의 자율주행자동차 사고는 보행자 사망으로 이어져 자율주행자동차의 윤리 문제에 대한 논의가 대두되었다.
- 이는 과거 트롤리 딜레마 이슈와 관련하여 지속적으로 논의되었으나, 자율주행자동차 사고들은 트롤리 딜레마 상황으로 발생한 것으로 보기에는 어렵다*는 의견이 모아지면서, 최근 현실적인 운행 환경에서 발생할 수 있는 위기 관리 이슈로 확산되고 있다.
 - * Jenkins, R., Goodall, N. J., Bonnefon J 등의 연구에서 트롤리 딜레마 상황으로 인해 자율주행자동차 사고가 발생하는 것은 아니라는 의견이 모아짐
- 이에 자율주행자동차 윤리 문제와 관련된 내용을 담은 지침을 세계최초(‘17.6)로 독일에서 ‘Ethics commission-automated and connected driving’를 발표하였으며, 미국에서도 ‘16년 9월 설계지침을 발표한 이후 매년 지속적으로 자율주행자동차 설계지침을 제시하고 있다.
- 국내에서도 자율주행자동차 행위주체들이 참고, 활용할 수 있도록 국내외 인공지능/자율주행자동차 가이드라인, 설문조사, 빅데이터 분석 등을 통해 본 자율주행자동차 윤리 가이드라인을 마련했다.

- 본 가이드라인을 통해 자율주행자동차의 기본가치와 행동원칙을 제시하고, 설계·제작·관리·서비스 제공·이용자 등 자율주행자동차 관련 행위주체들이 이를 고려하여 향후 상용화될 자율주행자동차가 우리 사회에 정립된 윤리 수준에 따라 제작·운행할 수 있도록 자발적으로 책임성을 확보할 수 있도록 한다.

2. 자율주행자동차 윤리 가이드라인 제정 경위

- 자율주행 상용화와 기술발전에 대비하기 위해 ‘자율주행자동차 차량·운전자 제어권 전환 안전성 평가기술 및 사회적 수용성 연구’ R&D 사업을 '17년부터 '21년까지 추진하였다.

* ‘자율주행자동차 윤리 가이드라인’ 연구는 홍익대 황기연 교수가 수행하였음

- '17년부터 국내외 관련 사례를 검토하고 전문가 의견 수렴을 통한 연구를 진행하였으며, ‘2019년 자율주행차 융복합 미래포럼 성과발표회’에서 자율주행자동차 윤리 가이드라인 초안을 발표하였다.

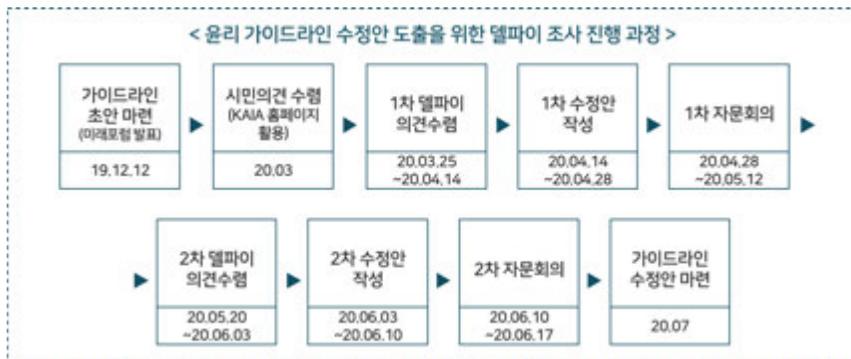
- 윤리가이드라인 초안은 다음 연구과정 진행 등을 통해 작성하였다.

- 1) 자율주행자동차 관련 국민 의식조사 1회
- 2) 윤리 가이드라인 가안에 대한 국민의견 수렴 2회,
- 3) 윤리 가이드라인 핵심요소 도출을 위한 11개 윤리 가이드라인 키워드 빅데이터 분석 수행



2019년 자율주행차 융복합 미래포럼 성과발표회('19.12.12)

- 초안 발표 이후 다양한 의견을 수렴하기 위해 국토교통과학기술진흥원 홈페이지에 초안을 공개하여 국민 및 유관 기관 의견을 수렴하였다.
- 또한 자율주행자동차 윤리 가이드라인의 객관성을 확보하기 위해 2차례 델파이 전문가 조사를 실시하여 수정안을 마련하였다.



자율주행자동차 윤리 가이드라인 수정안 도출을 위한 델파이 조사 진행 과정

- 델파이 조사를 통해 도출된 ‘자율주행자동차 윤리 가이드라인 수정안’은 ‘자율주행자동차 윤리 가이드라인 공개 토론회’에서 발표되었다.
- 윤리 가이드라인이 자율차 사고 배상책임 판단의 근거가 될 수 있으므로 행위주체별 책임 규정 시 신중을 기할 필요가 있다는 의견과 함께 업계의 기술개발을 촉진할 것으로 기대된다는 등 활발히 의견이 개진되었다.



자율주행자동차 윤리 가이드라인 공개토론회(‘20.08.14.)

- 공개 토론회에서 논의된 내용 등을 반영한 후 자동차 및 윤리 전문가 집단 등과 공동 제안 발표를 하였고, △자율주행자동차 기본가치, △자율주행자동차 행동원칙, △자율주행자동차 관련 행위주체(설계제작·관리·서비스제공·이용자)별 책임사항 등을 제시한 최종 가이드라인을 마련하였다.

- 공동제안에 참여한 기관 및 학회는 총 8개로 다음과 같다.
- 국토교통과학기술진흥원, 대한교통학회, 인공지능윤리협회, 한국교통안전공단, 한국교통연구원, 한국윤리학회, 한국자동차공학회, 한국자동차안전공학회



인공지능 윤리 대전 AI for Human('20.09.17.)

02

적용범위 및 구성체계

02 적용 범위 및 구성 체계

- 본 윤리 가이드라인의 적용대상은 정해진 조건에서 운전엔 전혀 개입하지 않거나, 사람이 직접 운전엔 관여하지 않는 레벨4 이상 완전 자율주행 시스템이 장착된 자동차를 설계·제작·관리·운영하려는 자와 사용자(소비자) 등을 대상으로 한다.
- 본 윤리 가이드라인은 ‘총칙’, ‘공통원칙’, ‘행위주체별 책임’으로 구성된다.
- ‘총칙’은 윤리 가이드라인의 제정 이유를 밝히는 ‘목적’, ‘용어의 정의’, 윤리 가이드라인의 ‘적용 대상’ 등으로 구성된다.
- ‘공통원칙’은 ‘기본가치’, ‘행동원칙’으로 구성된다.
- ‘행위주체별 책임’은 책임을 분담하는 행위주체인 ‘설계자’, ‘제작자’, ‘관리자’, ‘서비스 제공자’, ‘이용자’로 구성되며, 5개 행위주체에 대해 총 28개의 항목으로 구성한다.

03

총칙

1. 목적
2. 용어의 정의
3. 적용대상

03 총칙

1 목적

- 자율주행자동차는 과학·기술의 급속한 발전으로 인해 인간이 직접 운전하는 전통적인 자동차에 비해 운행 시 인간의 생명과 신체의 안전을 획기적으로 개선할 것으로 예상된다.
- 그러나 비록 자율주행자동차가 완성도 및 기술적 측면에서 고도로 발전하더라도 사고 등의 안전문제, 생명·재산 등의 손실, 그 밖에 미리 예상하기 어려운 문제들을 완벽하게 피하거나 방지하는 것은 현실적으로 불가능할 수 있다.
- 자율주행자동차가 자율주행시스템을 통해 사람에 의한 운전 개입 없이 상황을 판단하고 대응한다고 하더라도 위와 같은 문제들에 대하여 스스로 책임을 지기는 어렵기 때문에 결국 자율주행자동차의 설계자·제작자·관리자·서비스 제공자·이용자 등이 책임을 진다.
- 그러나 책임의 판단 기준을 사전에 세세히 정할 수 없기 때문에¹⁾, 사고로 인한 손실을 예방하고²⁾ 상용화 과정에서 윤리 문제로 인한 걸림돌을 사전에 제거하기³⁾ 위해서는 윤리 가이드라인이 필요하다.
- 본 자율주행자동차 윤리 가이드라인의 제정 목적은 자율주행자동차의 도입으로 편리하고 자유로운 이동권 보장 및 교통사고 피해 감소 등의 긍정적 효과를 증진시키고, 자율주행자동차의 도입으로 발생할 수 있는 부정적 효과를 감소시키는 동시에⁴⁾ 책임 소재를 가릴 수 있는 방향을

제시하며⁵⁾, 관련 행위주체들이 자율주행자동차의 기본 가치에⁶⁾ 따라 각각의 행동원칙을⁷⁾ 준수할 것을 권고하기⁸⁾ 위함이다.

- 본 윤리 가이드라인은 자율주행자동차에 관한 기본 지침적 성격을 가지며, 설계, 제작, 관리, 서비스 제공 및 이용 등 자율주행자동차와 관련된 전반적인 사항에 적용된다.

- 1) 자율주행자동차의 윤리 문제는 “자율주행자동차 운행 중 사고를 회피할 수 없는 상황을 접했을 때 법을 준수한 소수와 법을 어긴 다수 중 어느 쪽을 희생시킬 것인가”와 같은 트롤리 딜레마 상황을 비롯하여 신호등 고장, 차량사고 현장을 회피하여 통행을 진행하기 위해서는 불가피하게 법을 어기고 중앙선을 넘어야 하는 경우와 같은 위기관리 상황 등 다양하게 나타나기 때문에 상황별로 판단 기준을 관련법이나 규칙을 통해 사전적·명시적으로 규정하기는 어려움.
- 2) 트롤리 딜레마 상황과 같이 자율주행자동차 운행 시 피할 수 없는 사고가 발생하여 윤리적 판단을 내려야 하는 경우에는 자율주행자동차가 어떠한 선택을 하든지 필연적으로 인적·물적 피해를 발생시킬 가능성이 큼. 따라서 이러한 피할 수 없는 사고와 그에 따른 윤리적 선택 상황이 발생하지 않도록 예방 조치를 취하는 것이 중요함.
- 3) 본 윤리 가이드라인은 자율주행자동차 운행에 관련된 윤리적 명제에 대한 결론을 내리는 것보다는 상용화에 걸림돌이 될 수 있는 윤리적 문제의 방지를 위한 방향을 제시한다는 차원에서 제정 의의가 있음.
- 4) 긍정적 효과를 증진하고 부정적 효과를 감소시키는 것이 앞서 언급된 손실 발생의 예방과 함께 본 가이드라인이 공리주의적 원칙을 따르고 있다는 의미로 해석되기 때문에 ‘공리주의’ 원칙을 표면화하지는 주장도

있지만, 이 경우 교통약자나 사회적 약자들에게 불리한 형평성 문제를 초래할 가능성이 있어서 이를 반영하기는 어려움.

- 5) 자율주행자동차가 사물이어서 스스로 책임을 질 수 없기 때문에 사고 또는 문제 발생 시 자율주행자동차 관련 설계자, 제작자, 관리자, 서비스 제공자 및 이용자 등 행위 주체들이 책임의 주체가 됨.
- 6) 본 가이드라인의 적용 대상인 자율주행자동차는 주변상황과 도로 정보 등을 스스로 인지하고 판단하여 자동차를 운행할 수 있게 하는 자율주행시스템을 통해 작동되는 것으로, 인간의 개입 없이 자체적인 인지·판단을 통해 상황에 대응한다는 점에서 AI/로봇의 특성을 공유한다고 볼 수 있음.

이에 따라, 기본가치 제정 시 그동안 국내외 우수 기관의 ‘AI/로봇 윤리가이드라인’에서 제시된 관련 내용을 일부 반영하였음.

이제까지 공표된 국내외 AI/로봇 윤리 가이드라인(한국공학한림원, 2019)이 추구하는 기본가치의 핵심 내용은 ① 인간은 존엄하기 때문에 인간의 명령에 따라야 하고, ② 인간의 행복을 추구하여야 하며, ③ 인간의 생명을 보호하고 공공선을 추구하며, ④ 환경적, 사회적으로 지속가능한 발전을 추구해야 함 등임.

- 7) 자율주행자동차 행위주체들이 기본가치 구현을 위해 준수해야 할 행동원칙은 ① 투명성, ② 제어 가능성, ③ 책임성, ④ 안전성, ⑤ 보안성 등 크게 5가지로 구성됨.
- 8) 본 가이드라인은 행위주체들에게 의무사항을 강제하기 위한 것이 아니라, 우리 사회에 이미 정립된 윤리를 바탕으로 한 자율주행자동차의 기본가치, 행동원칙 등을 공개적으로 제시하여 행위주체들이 이를 고려하여 행동할 수 있도록 하고 자발적으로 책임성을 확보할 수 있도록 하려는 것임.

2 용어의 정의

- ‘자율주행자동차’란 운전자 또는 승객의 조작 없이 자동차 스스로 운행이 가능한 자동차를 말한다.
- ‘자율주행시스템’이란 운전자 또는 승객의 조작 없이 주변 상황과 도로 정보 등을 스스로 인지하고 판단하여 자동차를 운행할 수 있게 하는 자동화 장비, 소프트웨어 및 이와 관련한 일체의 장치를 말한다.⁹⁾

9) 자율주행자동차와 자율주행시스템의 정의는 각각 「자동차관리법」 제2조제1호의3과 「자율주행자동차 상용화 촉진 및 지원에 관한 법률」 제2조제1항제2호를 따름.

- ‘설계자’란 인공지능, 부품, 제어기, 인간-기계 상호작용 등 자율주행 자동차 기술의 하드웨어 또는 소프트웨어 분야에서 각종 이론, 요소 기술, 통합기술 등을 활용해 자율주행자동차를 설계하는 자 또는 단체, 기관, 회사를 말한다.
- ‘제작자’란 자율주행기술을 제작하거나, 다양한 자재와 기초 부품을 이용해 자율주행자동차를 최종적으로 조립 및 완성, 판매 및 공급하는 자 또는 단체, 기관, 회사를 말한다.
- ‘관리자’란 자율주행자동차의 안전한 운영을 확보하기 위해 관련 인프라, 정책, 법, 운행 허가, 사고 모니터링, 교육 등에 대한 관리의 책임과 의무를 지는 자 또는 정부(국토교통부 등), 그리고 정부로

부터 권한을 위임받은 자 또는 단체, 기관을 말한다.

- ‘서비스 제공자’란 제작된 자율주행자동차를 이용해 시장의 수요에 맞춰 서비스를 제공하고 운영하여 부가가치를 창출하는 자 또는 단체, 기관, 회사를 말한다.
- ‘이용자’란 자율주행자동차가 제공할 수 있는 서비스를 직·간접적으로 이용하는 자 또는 단체, 기관, 회사를 말한다.

3 적용 대상¹⁰⁾

- 본 윤리가이드라인에서 대상으로 하는 자율주행자동차는 운전자가 정해진 조건에서 운전 전 전혀 개입하지 않거나, 또는 모든 상황에 개입하지 않는 기술 수준에 해당된다.

10) 본 가이드라인의 적용 대상은 미국자동차공학회(SAE)의 자율주행 자동차 등급 기준에 따르면 Level 4 이상에 해당됨. Level 3 자율주행자동차는 비상상황 발생 시 운전자에게 제어권을 전환하거나, 위험 최소화 운영을 하는 등 능동적으로 대응하는 능력이 부족하기 때문에 윤리 가이드라인의 적용 대상으로 보기 어려움.

본 윤리 가이드라인에서 대상으로 하는 자율주행자동차는 운전자가 정해진 조건에서 운전 전 전혀 개입하지 않거나, 또는 모든 상황에 개입하지 않는 기술 수준에 해당되며, 독일 자율주행자동차 윤리강령에서도 적용 대상을 이같이 규정하고 있음.

04

공동원칙

1. 기본가치
2. 행동원칙

11) 공통 원칙을 제정한 이유는 현재까지 독일과 미국 등에서 발표된 자율주행자동차 관련 윤리 및 안전 가이드라인들에서 자율주행 자동차가 추구하는 기본 가치와 행위주체별로 준수해야 할 행동 원칙 등에 대한 구체적 언급이 없어서 각 행위 주체가 준수해야 할 의무 사항들이 포함되었는지 파악하기 어려웠기 때문임(황기연, 변순용, 2019).

이러한 문제를 극복하기 위해 본 가이드라인에서는 IEEE(2017), EU(2019), German Ethics Commission(2017), NHTSA(2016) 등 전 세계 11개 AI/로봇/자율주행자동차 윤리가이드라인과 국내 한국로봇학회(2018)와 한국공학한림원(2019) 등이 제안한 AI윤리 가이드라인을 참고함과 더불어 시민·전문가·학계·업계의 의견수렴을 통해 공통 원칙을 도출했음.

1 기본가치

1.1. 자율주행자동차는 인간의 존엄성, 국제법적으로 인정된 인권과 자유, 프라이버시 및 문화적 다양성을 존중하고, 인간을 성별, 나이, 인종, 장애 등을 이유로 차별하지 않으며, 인간의 법과 관습에 의한 판단과 통제에 따르도록¹²⁾ 설계, 제작, 관리되어야 한다.

12) AI/로봇윤리 가이드라인(IEEE, 2017; EU, 2019)에 따르면 인간이 존엄하다는 것은 단순히 자율주행시스템이 인간에게 해를 끼쳐서는 안된다는 의미를 넘어 인간의 판단과 통제를 따라야 하는 제어 가능성(Controllability)이 필수라는 것을 뜻함. 자율주행자동차는 사물이



기 때문에 윤리적 판단을 포함한 모든 판단은 이성을 가진 인간만이 할 수 있기 때문임.

하지만 악의를 가진 인간은 자율주행자동차를 사람이나 동물의 생명을 살상하거나 재물을 파괴하는 등 사회적 혼란을 초래할 목적으로 사용할 수 있음.

이러한 문제를 대비하기 위해 인간의 ‘올바른’ 판단과 통제만 따르도록 할 필요가 있으나, 자율주행자동차가 스스로 올바른 여부를 판별할 수는 없기 때문에 올바름에 대한 정의가 필요함.

올바름을 명확하게 정의하기는 어려우나, 결국 사회에 정립된 윤리와 가치판단을 바탕으로 할 것이라고 볼 수 있음.

따라서 ‘올바름’이란 최소한 사회 구성원들의 합의를 기초로 한 법과 관습을 따르는 것으로 볼 수 있음.

- 1.2. 자율주행자동차는 인간의 행복과 이익의 증진을 위한 수단으로서 인간의 안전하고 편리하며 자유로운 이동권을 보장하고, 타인의 권리와 자유를 침해하지 않도록 설계, 제작, 관리되어야 한다.¹³⁾

13) 본 항목은 ‘인간의 행복추구권’에 관한 내용으로 AI/로봇 윤리가이드 라인에서 숙의된 부분이 자율주행자동차에 적용된 것임. 자율주행 시스템에 의한 운전은 사람이 하는 경우에 비해 교통사고 발생을 현저하게 줄일 수 있을 것으로 전망되는데 그 이유는 일반 자동차 교통사고 대부분이 인간의 실수로 발생하기 때문임.

또한 자율주행자동차는 인간이 직접 운전할 필요가 없거나 아주 제한된 경우에만 운전을 하기 때문에 운전에 따른 피로나 스트레스가 줄어들고 운전시간을 타 용도로 쓸 수 있어 생산성 향상에도 기여하고, 고령, 장애인, 어린이 등과 같은 교통약자들의 이동권도 개선함.

하지만 자율주행자동차는 보행자, 이동 중인 차량, 도로의 형상 및 지형, 인터넷 및 주변 인프라 등 혼잡한 교통환경의 실제 도로를 주행할 것이기 때문에 사고의 위험이 있고, 트롤리 상황과 같은 예기치 못한 딜레마에 처할 수 있음.

피할 수 없는 사고 상황에 처한 자율주행자동차가 탑승한 승객을 보호하기 위해 고의적으로 사고를 내어 주변을 운행하는 차량이나 보행자들에게 피해가 가도록 설계되어 있다면 이는 타인의 이동의 자유를 침해하는 것으로 수용될 수 없음.

- 1.3. 자율주행자동차는 자동차 사고로 인해 발생할 손실을 최소화하고, 무엇보다 인간의 생명을 우선하도록¹⁴⁾ 설계, 제작, 관리되어야 한다. 또한, 손실을 최소화하는 과정에서 인간을 성별, 나이, 종교 등 개인적 차이 등을 이유로 차별하지 않고¹⁵⁾, 교통 약자를 고려하는 방식으로 작동하도록 설계, 제작, 관리되어야 한다.¹⁶⁾

14) 자동차 사고로 인해 발생할 수 있는 손실은 인간의 생명, 신체 등 인적 손실과 재산 등 물적 손실을 포함하는 의미임.

한편, 독일 자율주행자동차 윤리강령 7조에서는 “만약 인간이 다치는 것을 피하기 위해서라면 동물이나 물건을 다치도록 하는 것이 기술적으로 가능하도록 자율주행시스템을 설계해야한다”라고 규정하고 있음.

본 가이드라인을 도출하는 과정에서 다수의 시민과 전문가들은 “인간의 생명과 신체의 보호를 위해 동물의 생명 등을 무분별하게 희생시킬 수 있다”는 왜곡된 해석이 발생하지 않도록 해야 한다는 의견을 밝혔음. 어느 한쪽을 선택하라고 명시하는 순간 인간 생명의 보호라는 기본 가치와 충돌되기 때문에 이에 대해서는 추가적인 논의가 진행 중임.

15) 독일 자율주행자동차 윤리강령 제8조 ‘...인간 운전자가 위급한 상황에서 여러 명의 사람을 구하기 위해 한 사람을 죽인다면 잘못된 행동을 한 것이기는 하지만, 반드시 법적으로 유죄가 되는 것은 아니다...’를 근거로 일부 시민과 전문가들은 사람 인원수에 따른 차별도 명시화해야 한다는 주장이 있었음.

그렇지만 차량 내·외부의 인원수를 파악하는 기술적 문제와 인원수에 따른 차별을 금지할 경우 발생할 수 있는 또 다른 윤리 문제를 고려할 경우 인원수에 따른 차별 금지가 반드시 윤리적이라고 보기는 어렵다는 결론에 이르렀음.

또한, 법 준수 여부에 따라 차별할 수 있다는 주장도 제기되었음. 예를 들어 자율주행자동차가 피할 수 없는 선택 상황에서 법을 지켜서 안전모를 쓴 모터사이클 운전자와 법을 어기고 안전모를 쓰지 않은 모터사이클 운전자 중 충돌 방향을 선택해야 하는 경우, 안전모를 쓰지 않은 모터사이클 운전자와 충돌하면 사망할 가능성이 높기 때문에 안전모를 쓴 모터사이클 운전자와의 충돌을 선택할 수 있음.

이러한 문제에 대응하기 위해 “법을 위배하지 않는 범위 안에서 총체적인 피해를 최소화(legally-adjusted aggregate harm minimization)”해야 한다는 개선 방안이 제시되었지만 이 경우도 어느 한쪽을 선택하라고 명시하는 순간 인간 생명의 보호라는 기본 가치와 충돌되기 때문에 이에 대해서는 추가적인 논의가 진행 중임.

16) 교통약자를 고려하는 것이 인간을 차별하지 말라는 전 구절과 배치된다는 의견이 있지만, 구별과 차별은 다르기 때문에 큰 문제가 되지 않음. 같은 교통사고를 당하더라도 상대적으로 보다 큰 손상을 입을 수 있는 교통약자에게 더 많은 배려를 하는 것이 정의라는 롤스(John Rawls)의 MaxMin 원칙에 부합되기 때문임(Jenkins, 2016).

1.4. 자율주행자동차는 지구 온난화, 교통체증 등을 고려하여 지속 가능한 발전이 이루어지도록 설계, 제작, 관리, 운영되어야 한다.¹⁷⁾

17) 자율주행자동차 이용이 늘수록 버스기사, 택시기사, 택배기사, 화물차 운전자 등의 직업 안정성에 타격을 줄 수 있고, 대중교통 이용과 보행이 줄어 비만과 같은 건강 문제가 심각해질 수 있으며, 자율주행자동차가 기존의 자동차보다 더 많은 화석연료를 사용하도록 하는 방식으로 운행되도록 설계되는 경우 기후 온난화와 같은 환경적 위해를 줄 가능성도 있음.

따라서 자율주행자동차를 설계, 제작, 관리, 운영하는 과정에서 행위주체들이 이러한 문제들을 고려하는데 그치지 않고 예방 의무까지도 부담해야 한다는 일부 시민과 전문가들의 주장이 있었지만 적절한지에 대하여는 추가적인 논의가 필요한 상황임.

- 2.1. 자율주행자동차는 운행 중 발생할 수 있는 문제 상황에 대비하여 차량의 설계, 제작, 관리, 서비스 제공, 이용 등의 행위 주체별 책임을 판단할 수 있도록 운행 정보에 대한 투명하고 추적 가능한 기록시스템을 갖추어야 한다.¹⁸⁾ 그리고 기록된 정보는 차량 보유자가 보관하되¹⁹⁾ 사고대응 또는 사후처리를 위해 필요로 하는 관계자에게 제공될 수 있어야 한다.²⁰⁾

18) 인간이 자동차를 운전하다가 사고가 발생하면 운전자가 사고에 대한 책임을 지면되지만 자율주행자동차는 운전자의 개입 없이 자율주행시스템이 차량을 운행하기 때문에 사고가 발생했을 경우 사물에게 책임을 묻고 상응하는 배상을 요구하는 것이 불가능함.

투명성(Transparency)은 자율주행자동차의 안전한 운영을 보장하기 위해 행위주체 간 사고에 따른 책임을 구분해서 물을 수 있어야 한다는 것을 의미하고, 추적가능성(Traceability)은 자율주행자동차에 문제가 생기거나 사고가 발생했을 경우, 자율주행시스템의 판단에 대해 문제가 있었는지를 파악하는데 필요한 각종 데이터에 신속하게 접근할 수 있어야 한다는 것을 의미함.

한편, 일부 시민과 전문가들로부터 상황이나 사고에 대해 일반인들도 쉽게 이해할 수 있도록 단어나 문장, 그림, 데이터 등을 이용해 표현한다는 것을 의미하는 설명 가능성(Explainability)도 명시해야 한다는 주장이 있었지만, 광의의 투명성이 이를 포함하고, 전문적인 내용을 일반인들이 이해할 수 있도록 표현한다는 것이 적절한 지에 대한 이견이 있어서 추가적인 논의가 필요한 상황임.

19) 자율주행자동차 운행 중 기록된 정보는 차량 보유자에게 귀속되기 때문에 보관 의무 또한 함께 귀속됨.

20) 자율주행자동차 사고와 관련하여 자율주행시스템이 판단하는데 사용된 각종 센서 데이터, GPS 데이터, 정밀지도 데이터 등에 대해 제3자의 요청이 있을 때 공개되거나 공유되어서 사고 원인을 정확하게 밝혀낼 수 있어야 한다는 의미임.

- 2.2. 자율주행자동차는 필요시 운전자 또는 탑승자의 판단에 의해 제어 또는 정지될 수 있는 기능을 갖추고 있어야 하며, 자율주행자동차의 올바른 운영을 위해 정부 또는 법에 의해 권한을 위임받은 기관은 차량의 검사, 관리, 서비스 제공, 이용 과정 등을 관리·감독할 수 있다. 다만, 관리·감독 과정에서 프라이버시 침해 등의 부작용이 나타나지 않도록 주의하여야 한다.²¹⁾

21) 본 항목은 ‘제어 가능성’ 원칙을 자율주행자동차에 적용한 것임. ‘제어 가능성’이란 자율주행자동차가 운행 중 오작동으로 인해 도로교통 법규를 위반하는 등 이상 징후를 보일 때는 1차적으로 탑승자가 그 작동을 즉각적으로 제어 또는 정지시킬 수 있어야 하고, 만약 탑승자의 제어가 어려울 경우에는 정부로부터 권한을 위임받은 기관이 비상 상황에 대처해야 함을 의미함.

또한, 정부로부터 자율주행자동차의 이상 징후를 관리할 수 있는 권한을 위임 받은 기관은 설계부터 이용에 이르는 전 과정에서 오작동이 발생하지 않도록 예방하거나 긴급 통제하는 역할 뿐 아니라, 운행 과정에서 생성되는 개인 정보가 인가자의 허락 없이 유출되지 않도록 관리·감독하는 기능도 수행해야 함.

2.3. 자율주행자동차의 운행 과정에서 문제가 발생했을 경우, 차량의 설계, 제작, 관리, 서비스 제공, 이용 등의 행위 주체는 각각의 문제에 상응하는 책임을 진다.²²⁾

22) 본 항목에서 말하는 책임이란 범위를 구체화할 수 없는 공통의 책임 개념에 해당되는 Responsibility와는 달리 특정인이 구체적 책임을 지는 Accountability의 개념으로, 단순히 인적 물적 손해를 배상한다는 범위를 넘어, 배상의 정도를 정확하게 측정하기 위해 문제가 발생하게 된 이유를 객관적이고 상세하게 설명할 수 있어야 한다는 것을 의미함.

일반적으로 상품에 문제가 생길 경우 소비자가 책임지거나 제작자가 제조물에 대한 책임을 지지만 자율주행자동차는 자율주행시스템에 의해 운행되기 때문에 문제가 생길 경우 이러한 자율시스템을 작동 시키는데 참여한 모든 행위주체들이 각자의 몫에 해당되는 분산된 책임(distributed responsibility)의 주체가 되어야 하고 배상 책임도 분산됨.

한편, 지능형 로봇의 개발이 속도를 내면서 자율주행시스템이 일으키는 사고에 대해 민법의 법인처럼 인공적으로 만들어진 개체(ADSE: Autonomous Driving System Entity)에게 계약이나 소송 등의 법적 권리와 책임을 부여하려는 논의가 진행 중에 있음(자율주행차의 법과 윤리, 2020).

2.4. 자율주행자동차는 운행 중에 사고가 발생하거나 발생이 예상되는 경우, 인간의 안전을 최우선으로 보호하기 위한 사고 예방, 충돌회피 등의 다양한 기술과 서비스를 적용하도록 설계, 제작, 관리되어야 하며, 차량의 서비스 제공 및 이용 과정에서 기술과 서비스에 대해 관련 주체들을 대상으로 안내, 교육, 훈련 등이 실시되어야 한다.²³⁾

23) 자율주행자동차의 주행 환경은 지속적으로 변하기 때문에 안전을 확보하기 위해 발생할 수 있는 모든 위해를 예방하거나 회피하는 것은 사실상 어려움.

특히, 트롤리 상황과 같이 필연적으로 인간의 희생이 발생할 수밖에 없는 경우도 발생할 수 있기 때문에, 피해의 정도를 완화시키거나 빈도를 감축하는 방안, 탑승자와 비탑승자들에게 사고 상황을 잘 대응하게 하는 방안 및 사고 발생 이후 자율주행자동차를 조치하는 방안에 대해 현장 및 가상 실험을 통해 사전에 유효성을 검증해야 함.

또한, 자율주행자동차의 행위주체들이 각각의 분야에서 책임 있는 역할을 담당하기 위해서는 공신력 있는 기관이 제공하는 교육 및 훈련에 참가해야 함.

2.5. 자율주행자동차는 개인정보보호 및 사이버보안 규정을 준수하도록 설계, 제작, 관리되어야 한다.²⁴⁾

24) 자율주행자동차가 주행하기 위해서는 탑승자 개인을 포함해 각종 차량 센서를 통해 공급되는 차내 시스템 및 주변 환경에 대한 데이터를 수집해 활용해야 하는데 이러한 정보들은 개인들의 신상, 이동 경로, 시간 및 방문 장소 등에 대한 민감한 개인정보를 포함할 수 있기 때문에 정보 권리자의 사전 동의 없이 취득, 저장, 처리, 공유 및 배포되면 심각하게 개인의 프라이버시를 침해할 수 있음.

따라서 비식별 조치를 비롯한 철저한 보안으로 개인정보 침해의 문제가 발생하지 않도록 사전적 조치가 필요하며, 불법적인 해킹은 자율주행 시스템의 오작동 등 기능에 심대한 영향을 주어 사고를 유발할 가능성이 크기 때문에 이를 미연에 대비할 수 있는 사이버 보안 준칙을 마련해야 함.

05

행위주체별 책임

1. 설계자
2. 제작자
3. 관리자
4. 서비스 제공자 및 이용자

05 행위주체별 책임

1 설계자

- 3.1.1. 자율주행자동차 설계자는 운행 중 사고 발생 시 책임을 명확히 하기 위해 운행기록 등을 보관할 수 있도록 자율주행자동차를 설계해야 한다.²⁵⁾

25) 본 항목은 자율주행자동차 설계자의 투명성 준수 책임을 규정한 것임. 운행기록 및 제어권 전환 기록 등에 대한 보관 의무와 활용에 대한 사항은 모든 행위주체들에게 적용되기 때문에 행동원칙 2.1에서 기술하였음.

- 3.1.2. 자율주행자동차 설계자는 자동차의 운행과 관련된 제반 법규를 준수하도록 자율주행자동차를 설계해야 하며, 제반 법규가 적용되지 않는 상황에서의 사고 발생이 예상되는 경우 윤리 가이드라인의 기본 가치에 부합하도록 자율주행자동차의 작동에 대한 원칙을 설정하여야 한다.²⁶⁾

26) 제반 법규가 적용되지 않는 상황에서의 사고 발생이 예상되는 경우 설계자가 개별적인 상황 별로 판단 기준을 제공하는 것은 상황의 다양성을 감안할 때 현실적이지 않기 때문에 2장 1절의 기본 가치를 설계자의 판단 기준으로 활용하도록 하였음.

기본가치의 핵심 내용은 자율주행자동차는 ① 인간은 존엄하기 때문에 인간의 명령에 따라야 하고, ② 인간의 복리를 증진시켜야 하며, ③ 인간의 생명을 보호하고 공공선을 추구하고, ④ 환경적, 사회적으로 지속가능한 발전을 추구해야 함 등임.

한편 법규가 적용되지 않는 개별 상황별로 정부가 구체적 판단 지침을 제공해야 한다는 일부 시민의 주장이 있었지만, 논의 결과 현실적으로 정부가 판단지침을 세세히 주기 어렵다는 점에 시민과 전문가들 대부분이 동의함.

3.1.3. 자율주행자동차 설계자는 자율주행자동차가 안전하게 운행될 수 있도록 탑승자/자동차 및 차량 외부자/자동차 간의 상호작용을 고려해야 하며, 사고 시 탑승자 보호를 위한 시스템 안전 기능을 포함하여 설계해야 한다.

3.1.4. 자율주행자동차 설계자는 사전에 사고를 최대한 예방할 수 있도록 설계하여야 하며, 부득이하게 사고가 발생한 경우 탑승자를 비롯한 관련 당사자의 피해를 최소화하도록 설계하여야 한다.²⁷⁾

27) 자율주행자동차가 일단 트롤리 상황 등 딜레마에 처하게 되면 사고를 피할 수 없기 때문에 가능한 이러한 상황이 발생하지 않도록 사전에 예방하는 조치를 취하는 것이 대단히 중요함.

참고로 독일의 자율주행자동차 윤리 강령 5조 ‘자동화되고 네트워크화된 자율주행자동차 기술은 가능한 한 사고를 피해야 하고 난감한 상황이 절대 발생하지 않도록 디자인되어야 한다...’와 8조 ‘...자율주행자동차는 사고를 미연에 방지하도록 설계되어야 하지만...’에서도 예방설계의 중요성을 강조하고 있음.

3.1.5. 자율주행자동차 설계자는 자율주행자동차의 보유자나 이용자가 자율주행 시스템을 불법으로 개조하거나 임의로 변경하여 안전을 위해하는 행위를 방지할 수 있도록 설계하여야 한다.

3.1.6. 자율주행자동차 설계자는 해킹, 개인정보나 사생활 침해 및 자율주행 자동차를 대상으로 한 불법적 실험에 대한 대응 방안을 마련해야 하고, 설계의 오류로 인해 문제가 발생했을 경우 책임을 진다.

- 3.2.1. 자율주행자동차 제작자는 운행의 법적·윤리적 기준에 대한 투명성을 확보하기 위하여 운행 관련 내용을 기록하고 보관할 수 있도록 제작해야 한다.
- 3.2.2. 자율주행자동차 제작자는 관련 법규나 인증기준, 생명 윤리, 정보통신 윤리, 공학 윤리를 준수하여 자율주행자동차를 제작하고 판매하여야 한다.
- 3.2.3. 자율주행자동차 제작자는 자율주행자동차의 제작 상의 결함으로 인해 발생한 피해에 대하여 책임을 져야 한다.
- 3.2.4. 자율주행자동차 제작자는 자율주행자동차의 소유자 및 이용자에게 자율주행 기능 사용 시의 유의사항 및 운행 안전에 대한 설명을 제공할 의무를 가지며, 이를 문서로 작성해 제공해야 하고, 소유자와 이용자의 설명 요구에 대해 성실히 대답해야 한다.
- 3.2.5. 자율주행자동차 제작자는 자율주행자동차에 대한 사이버 보안시스템을 갖추고 항상 변경 내용을 제공해야 한다.

3 관리자

3.3.1. 자율주행자동차 관리자는 각각의 소관 사항에 따라 사고 발생에 따른 책임소재를 규명하기 위해 사고의 사후조치를 관리감독할 의무를 지며, 각 주체의 책임에 관한 제도나 조치사항을 마련하고 보완할 의무가 있다.

3.3.2. 자율주행자동차 관리자는 각각의 소관 사항에 따라 자율주행자동차가 가능한 한 사고를 피하고 난감한 상황에 처하지 않도록 관리하기 위해 도로 및 공간구조, 정보 통신 인프라 등의 물리적 인프라를 재정비하고, 자율주행 기능과 관련된 인력에 대한 홍보, 교육, 훈련, 교통 체계, 정비 서비스 체계 등을 포함한 사회적 인프라를 확충하기 위해 노력하여야 한다.²⁸⁾

28) 참고로 독일의 자율주행자동차 윤리강령 5조에서도 인간 보호를 위해 관리자의 물리적·사회적 인프라 확충을 통한 안전성 준수 책임을 명시하고 있음.

3.3.3. 자율주행자동차 관리자는 각각의 소관 사항에 따라 개인정보 보호를 위해 자율주행자동차 데이터 보관 및 폐기에 대한 지침을 이해관계자들에게 제공하여야 하며, 사이버 보안에 관한 제도나 조치사항을 마련하고 보완할 의무가 있다.

4 서비스 제공자 및 이용자

- 3.4.1. 자율주행자동차의 서비스 제공자는 이용자의 안전을 도모하고, 타인의 이익이나 공익을 침해하거나 위해를 가해서는 안되며, 자율주행 소프트웨어를 최신의 상태로 유지하여야 한다.²⁹⁾

29) 제작자가 안전성 개선을 위해 소프트웨어를 업데이트하더라도, 서비스 제공자가 이를 실행하지 않을 경우 탑승자와 차량외부자의 안전이 위협을 받고 더 나아가 공익을 침해하거나 위해를 가할 경우도 발생할 수 있으므로 서비스 제공자는 소프트웨어의 최신 상태 유지 의무를 준수해야 함.

- 3.4.2. 자율주행자동차의 서비스 제공자와 이용자는 정해진 목적과 기능에 따라 자율주행자동차를 운행해야 한다.³⁰⁾

30) 자율주행자동차 서비스 제공자 및 이용자는 행동 원칙 2.2에 명시된 것과 같이 사용과 관련된 법률 등에 따라 정해진 목적과 기능에 부합하도록 자율주행자동차를 운행해야 함.

- 3.4.3. 자율주행자동차의 서비스 제공자는 불법적 사용, 오남용 등을 최소화하기 위해 노력해야 하며, 자율주행자동차를 불법적으로 사용하거나 오남용한 이용자는 이로 인해 발생하는 문제에 대해 책임을 져야 한다.



3.4.4. 자율주행자동차의 서비스 제공자와 이용자는 자율주행자동차의 안전 문제를 일으킬 수 있는 임의 및 불법 개조³¹⁾를 해서는 안되며, 안전을 위한 충분한 교육을 이수해야 한다.

31) 본 항목은 자율주행자동차 서비스 제공자 및 이용자의 안전성 준수 책임을 규정한 항목으로 3.4.3의 서비스 제공자 및 이용자의 불법적 사용과 오남용으로 인한 책임성 준수 의무와는 차이가 있음.

3.4.5. 자율주행자동차의 서비스 제공자는 서비스 과정에서 생성되는 개인 정보를 보호해야 한다.

06

결론

06 결론

- 본 윤리 가이드라인은 자율주행자동차에 대한 사회적 수용성 제고를 목적으로 관련 행위주체들이 자율주행자동차의 기본 가치에 따라 각각의 행동원칙을 준수할 것을 권고하기 위해 마련하였다.
- 작성을 위해 IEEE의 Ethically Aligned Design Guideline, 독일의 자율주행자동차 윤리강령 등 전 세계 11개 인공지능 및 로봇 윤리 가이드라인을 참고하였고, 3차례의 시민의식 조사를 실시하였다.
- 또한 시민/기업/기관 의견 수렴, 2차례의 전문가 델파이 조사 및 자문, 대한교통학회, 인공지능윤리협회, 자동차공학회, 자동차안전학회, 한국윤리학회 등 주요 학회를 통한 심도 있는 의견을 반영하여 최종 안을 완성하였다.
- 본 가이드라인은 자율주행자동차의 도입으로 편리하고 자유로운 이동권 보장 및 교통사고 피해 감소 등의 긍정적인 효과를 증진시키고, 부정적 효과를 감소시킴과 동시에 운행 사고에 대한 책임소재를 가릴 수 있는 방향을 제시하여 자율주행자동차의 수용성을 높이고 상용화를 촉진시킬 것으로 기대된다.
- 본 윤리 가이드라인은 향후 자율주행 기술 발전, 상용화, 새로운 정책과 법 제개정 등으로 인한 변동사항에 따라 수정될 수 있으며, 자율주행 자동차 제작자, 연구기관 및 사용자(소비자)와 긴밀한 소통을 통해 모든 이해당사자의 의견이 충분히 반영될 수 있도록 할 것이다.



[참고 문헌]

- 이중기, 황기연, 황창근, 자율주행차의 법과 윤리, 박영사, 2020.
- 한국로봇학회, 지능형 로봇의 개발과 활용에 대한 윤리 가이드라인(안), 2018.10.31.
- 한국공학한림원, 한국형 AI 및 로봇 윤리가이드라인, 2019a.09.30.
- 황기연, 변순용 “자율주행자동차 윤리적 쟁점 검토 및 윤리 가이드라인 초안” 자율주행자동차 융복합 미래포럼 발표자료, 2019b.12.12.
- European Commission, Ethics Guidelines for Trustworthy AI, High-Level Expert Group on Artificial Intelligence, April 8, 2019
- German Federal Ministry of Transport and Digital Infrastructure, Ethics Commission: Automated and Connected Driving, June 2017.
- IEEE, Ethically Aligned Design-Version II, 2017.
- Jenkins, R. Autonomous Vehicles Ethics & Law: Toward an Overlapping Consensus, New America, September 2016
- National Highway Traffic Safety Administration. Federal automated vehicles policy: Accelerating the next revolution in roadway safety. US Department of Transportation. 2016 September.
- Goodall, N. J. (2016). Can you program ethics into a self-driving car? IEEE Spectrum, 53(6), 28-58.
- Bonnefon J. et.al.(2015) “Autonomous Vehicles Need Experimental Ethics: Are We Ready for Utilitarian Cars?”, arXiv:1510.03346
- Goodall, Noah J. (2014), “Ethical Decision Making During Automated Vehicle Crashes”, Journal of the Transportation Research Board, Transportation Research Board of the National Academies, 2424, pp.58-65.

06

부록

부록 1 자율주행자동차 윤리 가이드라인 관련 국내외 사례

1 Asimov의 로봇윤리 3원칙

- 아시모프가 최초로 제시한 로봇 윤리원칙 다음과 같음

원칙 0: 로봇은 인류에게 해를 끼쳐서는 안 되며, 위험에 처한 인류를 방관해서도 안 된다(A robot may not injure humanity, or, through inaction, allow humanity to come to harm).

원칙 1: 로봇은 인간을 다치게 해서는 안 되고, 또는 위험에 처한 인간을 방관해서도 안 된다(A robot may not injure a human being or, through inaction, allow a human being to come to harm).

원칙 2: 로봇은 인간이 내린 명령에 복종해야 한다. 다만 명령이 원칙 1과 상충되는 경우는 예외로 한다(A robot must obey the orders given to it by human beings, except where such orders would conflict with the First Law).

원칙 3: 로봇은 원칙 1, 2와 갈등하지 않는 한에서 자기를 보호해야 한다(A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws).

- 해당 로봇윤리 원칙에 대해서는 다양한 수정안들이 제안되었으나, 변하지 않는 원칙은 로봇은 어떠한 경우에도 인간의 존엄성을 최우선적으로 보호해야 한다는 것임

- 2017년 1월, AI 비영리 단체에서는 아실로마 AI 원칙을 발표하였으며, 윤리와 가치 부문에서 인공지능을 개발하는 과정에서 준수해야 할 13가지 원칙을 제시함

1. 안전 : 인공지능 시스템은 작동 수명 전반에 걸쳐 안전하고 또 안전해야 하며, 적용 가능하고 실현 가능할 경우 그 안전을 검증할 수 있어야 함.
2. 장애 투명성 : 인공지능 시스템이 손상을 일으킬 경우 그 이유를 확인할 수 있어야 함.
3. 사법적 투명성 : 사법제도 결정에 있어 자율시스템이 사용된다면, 권위 있는 인권기구가 감사 할 경우 만족스러운 설명을 제공할 수 있어야 함.
4. 책임 : 고급 인공지능 시스템의 디자이너와 설계자는 인공지능의 사용, 오용 및 행동의 도덕적 영향에 관한 이해관계자이며, 이에 따라 그 영향을 형성하는 책임과 기회를 가짐.
5. 가치관 정렬 : 고도로 자율적인 인공지능 시스템은 작동하는 동안 그의 목표와 행동이 인간의 가치와 일치하도록 설계되어야 함.
6. 인간의 가치 : 인공지능 시스템은 인간의 존엄성, 권리, 자유 및 문화적 다양성의 이상에 적합하도록 설계되어 운용되어야 함.
7. 개인정보 보호 : 인공지능시스템의 데이터를 분석 및 활용능력의 전제하에, 사람들은 자신들이 생산한 데이터를 액세스, 관리 및 통제할 수 있는 권리를 가져야 함.
8. 자유와 개인정보 : 개인정보에 관한 인공지능의 쓰임이 사람들의 실제 또는 인지된 자유를 부당하게 축소해서는 안됨.

9. 공동이익 : 인공지능 기술은 최대한 많은 사람에게 혜택을 주고 힘을 실어주어야 함.
10. 공동번영 : AI에 의해 이루어진 경제적 번영은 인류의 모든 혜택을 위해 널리 공유되어야 함.
11. 인간의 통제력 : 인간이 선택한 목표를 달성하기 위해 인간은 의사결정을 인공지능 시스템에 위임하는 방법 및 여부를 선택해야 함.
12. 비파괴 : 고도화된 인공지능 시스템의 통제로 주어진 능력은 건강한 사회가 지향하는 사회적 및 시정 과정을 뒤엎는 것이 아니라 그 과정을 존중하고 개선해야 함.
13. 인공지능 무기 경쟁 : 치명적인 인공지능 무기의 군비 경쟁은 피해야 함.

- 행동준칙: 투명성, 제어가능성, 책임성, 안전성, 보안성 강조
- 기본가치: 인간 존엄성, 권리, 자유, 문화적 다양성, 이익과 번영의 공유, 살상무기화 금지 강조

3 EU Ethics Guidelines for Trustworthy AI

- 2019년 EU에서 공포한 “신뢰할 수 있는 AI 윤리 가이드라인”은 기존의 로봇/AI 윤리 논의의 결론적 위상을 가짐
- 기본가치: 인간자율성, 위해 방지, 공정, 설명가능성
- 행동준칙: 제어가능성, 안전성, 보안성, 인간 복리, 비차별성, 책임성, 투명성

4 IEEE EAD(Ethically Aligned Design) Guideline

- IEEE의 지능형 자율시스템 윤리적 디자인(EAD) 가이드라인(2017)
- 국제전기전자공학협회(IEEE)의 윤리적 디자인(Ethically Aligned Design, EAD) 가이드라인의 원칙은 다음과 같음

- 1) 국제법적으로 인정된 인권, 자유, 인간 존엄성, 프라이버시 및 문화적 다양성을 침해해서는 안 되며, 보안과 안전 차원에서 검증이 가능하면서 인간에게 이롭도록 만들어져야 하고, 항상 인간의 판단과 통제에 따라야 함.
- 2) 디자인과 사용에 있어서 인간의 복리(Wellbeing)를 최우선 하여야 하며, 복리는 단순히 삶의 만족도 뿐 아니라 삶에 영향을 주는 사회 및 자연 등 다양한 조건의 충족도 동시에 긍정적인 효과와 부정적 효과의 적절한 균형까지 포함하는 개념임.
- 3) 설계자, 제조자, 사용자 및 운영자들은 문제가 발생할 경우 물질적 보상을 포함한 책임을 부담해야 하며, 법적 책임을 명확하게 하기 위해 등록, 활동 등에 대한 기록시스템을 갖추어야 함.
- 4) 어떻게, 왜 특정한 결정이나 행위를 했는지 추적, 설명 및 해석이 가능하도록 투명한 방식으로 설계, 제작, 사용되어 문제가 생겼을 때 이해관계자 별로 책임을 분배할 수 있고, 사회적 수용성을 높일 수 있어야 함.
- 5) 해킹, 프라이버스 침해, 과용 등과 같은 위험요인, 부정적 효과를 최소화해야 함.

- 기본가치: 인간의 존엄성, 인권, 자유, 문화적 다양성, 프라이버시 보호, 사회 및 환경적 복리, 인간에게 이롭게
- 행동준칙: 투명성(검증가능, 기록시스템, 추적/설명/해석 가능성), 제어가능성, 책임성(설계, 제작, 이용자간 책임분배), 보안성(해킹, 프라이버시 침해, 과용), 안전성
- 행위주체: 설계자, 제작자, 사용자, 관리자

5 Beijing AI 윤리 가이드라인

- 2019년 6월 베이징 BAAI(Beijing Academy of Artificial Intelligence)에서 “베이징 AI 원칙(Beijing AI Principles)”를 공표
- 연구개발, 사용자, 거버넌스의 대항목으로 구분하여 세부 요소 도출 및 가이드라인의 내용이 전반적으로 포괄적임
- 연구개발 분야는 총 7개의 항목으로 목표와 가치, 책임, 데이터의 보안 및 투명성을 중점으로 다루고 있으며 다수의 이익을 중요시 하는 것으로 확인
- 이용자 분야는 3개 항목으로 사용의 책임, 정보 동의, 교육 훈련에 대한 내용이 포함되며 이익 최대화, 위험성(Risk) 최소화를 강조
- AI 거버넌스 분야는 총 5개의 항목으로 고용 최적화, AI 생태계를 위한 공생, 규제 개정, 지침의 세분화, 장기적 전략을 세부 내용으로 다루며 시대 흐름의 변화를 잘 반영할 수 있는 전략적 설계 및 협력을 중시
- 기본가치: 인간복리(고용, 조화와 협력, 적응과 완화, 장기계획), 다양성과 포용성
- 행동준칙: 책임성, 안전성(이용자지침, 정보제공, 교육/훈련), 윤리성, 공개와 공유
- 행위주체: 설계자, 이용자, 관리자

6 독일 자율주행자동차 윤리 강령

- 지난 2017년 6월, 독일에서는 세계 최초로 자율주행자동차의 윤리 강령을 발표하였음
- 해당 강령은 독일 연방 교통부장관, 연방헌법재판소, 본 대학(Bonn University) 교수진을 중심으로 다양한 한계 전문가로 구성된 윤리 위원회(The Ethics Commission)에서 초안을 작성하였으며, 20개의 핵심 항목으로 구성되어 있음
- 다음의 5개 분야로 구분하여 윤리 가이드라인을 제시하고 있고, 각 항목별 윤리적 속성은 아래와 같음
- 독일 가이드라인은 윤리 핵심요소를 대부분 수용하고 있으며 특히, 차량디자인과 인간의 안전과 복리 증진을 강조함

1. 부분 혹은 완전 자율주행자동차는 우선적으로 도로 이용자들의 안전을 개선하고, 더 나아가 이동 기회 증진과 이를 통해 더 많은 이득을 얻도록 하는데 사용될 수 있다. 기술 개발은 사적 자치의 원칙에 따르지만, 개발자는 자신이 개발한 기술에 대해 책임을 져야 한다.
2. 인간 보호는 부분 또는 완전 자율주행자동차가 가져다주는 모든 효용 중 가장 중요한 것이다. 인간 보호란 단순히 피해를 줄이는 것이 아니라 완전하게 피해를 회피하는 것을 의미한다. 따라서 자율주행자동차를 사회적으로 수용하기 위해서는 동일한 위험의 수준 하에서 자율주행자동차가 인간이 운전할 경우에 비해 피해를 줄일 수 있다는 결과를 입증할 수 있어야 한다.

3. 대중이 이용하는 교통시스템에 적용되는 자율주행자동차에 대한 운행 허가는 공공기관이 담당하고 실제 운영을 할 때도 공공의 승인과 통제가 필요하다. 사고 회피는 지켜야 할 가장 중요한 원칙이지만, 자율주행자동차 도입에 따른 위험에 대해 아무리 철저하고 확실하게 대비했다 하더라도 기술적으로 완벽하게 피할 수는 없는 위험이 존재한다는 점은 고려되어야 한다.
4. 인간이 스스로 책임지는 결정을 할 수 있다는 것은 즉 각 개인이 자기 개발과 보호에 대한 요구를 할 수 있는 사회를 뜻한다. 모든 국가와 정치적 질서는 개인의 자유로운 발전과 보호를 지향해야 한다. 하지만, 자유주의 사회에서 기술을 개발할 때 지켜야 할 원칙은 개인의 자유로운 의사결정에 따르더라도 다른 사람들의 자유 및 안전에 손상을 끼쳐서는 안 된다는 것이다.
5. 자동화되고 네트워크화된 자율주행자동차 기술은 가능한 한 사고를 피해야 하고 난감한 상황이 절대 발생하지 않도록 디자인 되어야 한다. 난감한 상황이란 자율주행자동차량이 어떠한 선택을 하더라도 피해를 볼 수밖에 없는 딜레마 상황을 의미한다. 이를 위해 통제가 필요한 통행구역에 대한 경계 설정, 차량 센서 및 제동 성능, 위험 상황에 처한 사람을 위한 신호, “지능형” 도로 인프라를 통한 위험 예방 등 기존의 동원 가능한 모든 제어 기술들이 동원되어야 할 뿐 아니라 지속적으로 새로운 기술도 개발해야 한다. 획기적으로 교통안전을 개선할 수 있도록 차량 소프트웨어에 대해서는 방어적이고, 미래 지향적이며, 교통약자들을 위해 안전 운전을 유도하는 방향으로 개발 및 규제의 목표가 설정되어야 한다.

- 
6. 자동충돌회피 기능을 보유한 고도로 자동화 된 자동차의 도입으로 인해 충돌을 줄일 수 있다면 이 자동차는 사회적으로나 윤리적으로 수용이 가능하다. 반대로 완전 자율주행자동차의 사용을 법적으로 의무화하거나 실질적으로 부족한 기술의 사용을 강요하는 것은 윤리적으로 문제가 된다. 특히 그런 의무나 강요가 기술적으로 반드시 준수해야 하는 사항과 연관되어 있는 경우는 더욱 문제가 된다.
 7. 모든 기술적 예방 조치에도 불구하고 피할 수 없는 위험한 상황에서는 다른 어떤 법적 가치가 있는 대상 중에서도 인간의 생명을 보호하는 것이 가장 중요하다. 이러한 이유로 자율주행 자동차는 만약 인간이 다치는 것을 피하기 위해서라면 동물이나 물건을 다치도록 하는 것이 기술적으로 가능하도록 프로그래밍 되어야 한다.
 8. 선택 상황 하에서 어떠한 생명을 살릴 것인가 하는 것은 “예측할 수 없는” 사람들의 행동까지 혼합된 실제 상황에서는 난감하다. 그러한 결정 상황에 대해서는 명확하게 규정하기도 어렵고 윤리적 문제를 명확하게 하지 않은 채 대응 프로그래밍을 짤 수도 없다. 자율주행자동차는 사고를 미연에 방지하도록 설계되어야 하지만, 결과에 대한 난해하고도 직관적인 예측을 통해 표준화된 설계를 하기는 어렵다. 따라서 자율주행자동차는 경우에 따라서는 도덕적이고 책임감 있는 차량 운전자의 결정으로 대체하거나 또는 활용할 수도 있다. 인간 운전자가 위급한 상황 하에서 여러 명의 사람을 구하기 위해 한 사람을 죽인다면 잘못된 행동을

한 것이기는 하지만, 반드시 법적으로 유죄가 되는 것은 아니다. 법적 판단은 일이 벌어지고 나서 구체적 사안에 대해 이루어 지는데, 이를 추상적이고 일반적인 방법을 통해 일일이 사전에 예측해 판단을 내리고 프로그래밍까지 하는 것은 불가능하다. 따라서 이러한 난감한 상황에 대한 판단은 독립된 공공 기관(자동화 된 교통 시스템 조사를 위한 연방 기관 또는 자율화 및 네트워크화 된 보안을 위한 연방 사무소)을 통해 체계적으로 사안을 처리하는 것이 바람직하다.

9. 피할 수 없는 사고 상황의 경우, 개인의 특성(나이, 성별, 신체적 혹은 정신적 상태)을 고려해 차별하는 것은 엄격히 금지된다. 또한 특정한 사람을 살리기 위해 다른 사람을 희생시키는 행위도 금지한다. 부상자 수를 줄이기 위한 일반적인 프로그래밍은 정당화 될 수 있다. 한편, 이동활동에 참여함으로써 인해 차량 사고로 인한 위험에 노출된 사람이 보행자와 같은 비참 여자를 희생시켜서는 안 된다.
10. 자율주행자동차의 경우, 인간이 운전하지 않기 때문에 그 책임이 기술 시스템의 제조자와 운영자로, 그리고 사회기반시설 담당 및 정치적, 법적 의사결정 기관 등으로 분산된다. 법적 책임 규정을 만들거나 사법적 의사결정을 할 때 이러한 책임의 분산을 충분히 고려해야 한다.
11. 기존 제조물 책임과 동일한 원칙이 자율주행자동차에 의해 발생한 손상 책임에도 적용된다. 따라서 제조자와 운영자는



지속적으로 자율주행시스템을 최적화하고, 기술적으로 가능하고 합리적 방법으로 모니터링하고 개선해야 할 의무가 있다.

12. 일반 시민들은 신기술 및 그 사용법에 대해 충분히 개인 별 특성을 고려한 차별화된 교육을 받을 권리가 있다. 자율주행 자동차의 사용 및 프로그래밍에 대한 지침은 가능한 한 투명하고, 공개되어야 하며, 검증 능력이 있는 독립된 기관에서 점검해야 한다.
13. 디지털화된 교통인프라를 통해 모든 차량을 완벽하게 네트워크로 연결해서 중앙에서 통제하는 것이 철도나 항공처럼 가능하고 또한 의미가 있을 지 판단하기 어렵다. 교통 상황을 완벽하게 파악하는 것이 사실 상 불가능하고, 차량제어시스템의 조작도 완전하게 막을 수 없기 때문에 디지털화된 교통인프라를 통해 모든 차량을 완벽하게 연결해 중앙에서 통제하는 것은 윤리적으로 수용하기 어렵다.
14. 자율주행자동차의 운행이 정당화되기 위해서는 IT 시스템 조작이나 내부 시스템의 결함 등으로 자율주행자동차 사고가 발생했을 그에 따른 피해가 도로교통에서 완전히 신뢰를 잃을 정도로 심각한 수준이어서는 곤란하다.
15. 자율주행자동차가 주행하면서 생산하는 데이터 기반의 비즈니스 모델은, 데이터가 차량을 제어하는데 요긴하게 사용되는 지 여부에 상관없이, 데이터에 대한 자치권이 개인들에게

속해 있고 개인 정보는 보호되어야 하는 등의 이유로 활용에 제약이 따른다. 차량소유자나 이용자는 일반적으로 자신들의 차량 데이터를 팔거나 사용하는 결정을 한다. 개인들이 각자의 결정으로 그런 데이터를 판매할 수 있다는 것은 데이터가 또 다른 용도로 유용하게 쓰일 실용적 가치가 있다는 것을 의미한다. 검색 엔진이나 소셜 네트워크 운영자들이 이러한 개인들의 차량 데이터를 당연한 것처럼 마음대로 가져다 쓰지 못하도록 초기 단계에 대응해야 한다.

16. 무인운전시스템이 사용되었는지 또는 제어권 전환으로 인해 책임이 인간운전자에 있는지 여부를 명확하게 구분할 수 있어야 한다. 완전한 무인운전이 아닌 경우, 인간과 기계의 관계를 설계할 때 어느 쪽이 어떤 책임을 맡고 있는지, 그리고 제어권이 어디에 있는지 여부를 항상 명확하게 해야 한다. 제어권에 관련해서는 전환이 일어나 시간과 전환 법칙 위반에 대한 책임이 인간과 기계 중 어디에 있는지를 문서화해 저장해야 하며, 이는 특히 인간과 기계 사이에 제어권 전환이 어떻게 이루어지는지를 명확하게 하기 위해 반드시 필요하다. 자율주행 자동차 기술이 국경을 넘어 확산될 경우에 대비해 프로토콜 또는 문서화 의무사항에 대한 국가 간 호환성이 확보되어야 하고, 제품 인도 절차와 문서화(기록)에 대한 국제 표준화도 필요하다.
17. 고도화된 자율주행자동차의 소프트웨어와 기술은 비상사태로 인해 인간운전자에게 갑작스러운 제어권 이전이 거의 발생하지



않도록 설계되어야 한다. 또한 무리함이 없이 인간과 기계 사이의 효율적이며 신뢰할 수 있고 안전한 소통을 가능하게 하기 위해, 자율주행시스템은 높은 수준의 적응 기술을 활용하기 보다는 사람의 의사소통 행태에 잘 적응하도록 설계되어야 한다.

18. 기계학습이나 인공지능 기술이 중앙 시나리오 데이터베이스와 연결되어 자율주행자동차의 안전을 개선한다면 윤리적으로 허용된다. 인공지능 기술은 차량 제어 관련 기능에 대한 안전 요건을 충족하고 명시된 규칙을 준수하는 경우에만 활용이 가능하다. 비작동 시간을 포함한 인공지능의 일반사용설명서를 만들기 위해 관련 시나리오를 중앙시나리오 목록에 추가할 수 있다.
19. 자율주행자동차는 인간의 개입 없이 비상 상황을 스스로 대처해 안전한 상태로 변환할 수 있어야 한다. 이를 위해 안전 상태가 어떤 것인지, 또한 일반적으로 용인될 수 있는 변환의 방식은 무엇인지에 대한 표준화된 정의가 필요하다.
20. 자율주행자동차를 올바르게 사용하기 위해 교육이 필요하다. 또한 자율주행자동차를 올바르게 다루는 방법에 대해 운전자 훈련 과정에서 적절히 조정되고 실험되어야 한다.

7 미국 NHTSA 자율주행자동차 설계 지침

- 자율자동차 디자인 모니터링을 위한 15개 항목
- 제조사나 기타 관련 기업은 차량이나 장비의 설계 및 개발 시 공통으로 차량 데이터 기록 및 공유 기능을 갖추고, 안전 및 사이버 보안에 대응할 수 있어야 함
- 또한 HMI를 고려한 설계와 적합한 내충격성/탑승자 보호 설계를 차량에 적용해야 하며 소비자 교육이 제대로 이루어질 수 있도록 관리하여야 함
- 자율주행자동차의 개발 목표 간 충돌, 목표 내부의 충돌 등 딜레마 상황에 대처할 수 있는 윤리적 원칙의 개발이 필요
- 공통 적용 분야 외에도 제조사나 기타 관련 기업은 자율주행시스템에 해당하는 운행설계영역(ODD), 객체 및 사고상황 인지 및 대응(OEDR) 기능을 갖추고 고장 시 조치를 통해 최소 위험 상태로 전환하는 절차와 안전성을 보장하기 위한 시험 및 검증 방법을 개발해야 함

1. 데이터 기록 및 공유: 제조사 및 기타 관련 업체는 문제의 원인을 규명할 수 있도록 사건, 사고 및 충돌 데이터의 시험, 검증 및 수집을 위한 문서화된 프로세스를 갖춰야 한다.
2. 개인정보보호: 제조사 및 기타 관련 업체는 개별 또는 공동으로 소비자 개인정보 보호를 위한 조치를 취해야 한다.
3. 시스템 안전: 안전 위험이 없는 자율주행시스템을 위해서는 차량이 어떠한 종류의 기능적 오류에도 안전한 상태에 있도록 하기 위한 설계가 필요하며, 인공지능(AI), 머신러닝, 기타 관련 소프트웨어 알고리즘을 모니터링하여 유효성과 안전성을 확보해야 한다.

- 
4. 차량 사이버보안: 제조사 및 기타 관련 업체는 사이버보안 위험을 최소화해야 한다.
 5. HMI (Human Machine Interface): SAE Level 3 시스템의 경우 자율주행시스템은 인간운전자, 운전자, 탑승자, 외부액터(타차량, 보행자 등) 등과의 원활한 상호작용을 고려해야 한다.
 6. 내충격성: 자율주행자동차는 다른 차량과 충돌가능성에 대비해 첨단 센싱 기술로부터 얻은 정보를 이용하여 모든 연령과 체격 조건의 탑승자를 보다 안전하게 보호할 수 있는 탑승자 보호 시스템을 갖춰야 한다.
 7. 소비자 교육: 제조사 및 기타 관련 업체는 직원, 딜러, 유통업자와 소비자를 대상으로 한 교육 프로그램을 개발하고 문서화해야 한다.
 8. 등록 및 인증: 차량의 생애주기 동안 소프트웨어 업데이트로 인해 차량의 자동화 수준이 변경될 수 있기 때문에 제조사 및 기타 관련 업체는 주요 기능에 관한 변경된 정보를 차량의 운전자와 소유주에게 손쉽게 전달하기 위한 차량 내 수단을 제공해야 한다.
 9. 충돌 후 거동: 제조사 및 기타 관련 업체는 자율주행자동차가 충돌 사고 후 재가동되는 방법을 평가하고 시험하고 검증하기 위한 문서화된 프로세스를 갖춰야 한다.
 10. 연방법, 주법, 지역법: 제조사 및 기타 관련 업체는 관련 연방법, 주법 및 지방법 등 관련법을 준수하고 운용 지역의 도로 규정을 준수해야 한다.

11. 윤리적 고려사항: 자율주행자동차의 컴퓨터 "운전자"가 미리 프로그래밍된 결정 규칙 또는 머신러닝 절차에 따라 내리는 다양한 결정들은 윤리적 관점이나 함의를 갖는다.
12. 운행설계영역(ODD): 제조사나 기타 관련 기업은 자율주행자동차의 운행설계영역(ODD: Operational Design Domain)을 문서화해서 안전하게 운행될 수 있는 조건과 시스템 기능의 평가, 시험 및 검증을 위한 프로세스 및 절차를 마련해야 한다. 정의된 ODD를 벗어나는 경우, 해당 차량은 최소 위험 상태로 전환되어야 한다.
13. 객체 및 사고상황 인지 및 대응(OEDR): 객체 및 사고상황 인지 및 대응(Object and Event Detection and Response, OEDR)은 안전한 운행에 영향을 줄 수 있는 정상 주행 상황 뿐 아니라 특수한 상황(경찰의 교통정리, 공사 인부의 교통 통제, 긴급 대응반 등)을 포함한 다양한 상황에 대응할 수 있어야 한다.
14. 고장 시 조치: 제조사 및 기타 관련 업체는 문제 발생 시 인간 운전자에게 알림으로써 운전자가 차량 제어를 넘겨받거나 또는 자율주행시스템이 독립적으로 최소 위험 상태로의 전환을 위한 문서화된 프로세스를 마련해야 한다.
15. 검증 방법: 서로 다른 자동화 기능에 대한 범위, 기술 및 기능이 매우 다양하게 나타난다는 점을 고려하여 제조사 및 기타 관련 업체는 자율주행자동차의 높은 안전성을 보장하기 위한 시험 및 검증 방법을 개발해야 한다.

부록 2 델파이 참여 전문가 명단

· 1차 델파이 전문가 의견수렴 명단

NO	성명	소속	비고
1	문영준	한국교통연구원 교통기술연구소 소장	융복합 포럼 기술분과 위원
2	문희창	언맨드솔루션 대표	융복합 포럼 기술분과 위원
3	윤일수	아주대학교 교수	융복합 포럼 기술분과 위원장
4	정재욱	쌍용자동차 수석	융복합 포럼 기술분과 위원
5	정재일	국민대학교 교수	융복합 포럼 기술분과 위원
6	기승도	보험연구원 수석연구원	융복합 포럼 산업분과 위원
7	김범준	LG경제연구원 책임연구원	융복합 포럼 산업분과 위원
8	남두희	한성대학교 교수	융복합 포럼 산업분과 위원
9	박요한	삼성교통안전문화연구소 수석연구원	융복합 포럼 산업분과 위원
10	임현준	홍익대학교 교수	융복합 포럼 산업분과 위원장
11	정재윤	경희대학교 교수	융복합 포럼 산업분과 위원
12	허정윤	국민대학교 교수	융복합 포럼 산업분과 위원
13	김정룡	한양대학교 교수	융복합 포럼 생활분과 위원장
14	박용훈	교통문화운동본부 대표	융복합 포럼 생활분과 위원
15	정원섭	경남대학교 교수	융복합 포럼 생활분과 위원
16	차두원	한국과학기술기획평가원 연구위원	융복합 포럼 생활분과 위원
17	김을겸	한국자동차산업협회 상무	융복합 포럼 제도분과 위원
18	윤경한	법률사무소	융복합 포럼 제도분과 위원
19	윤영한	한국기술교육대학교 교수	융복합 포럼 제도분과 위원장
20	이경수	서울대학교 교수	융복합 포럼 제도분과 위원
21	이중기	홍익대학교 교수	융복합 포럼 제도분과 위원
22	이항구	한국산업연구원 선임연구위원	융복합 포럼 제도분과 위원
23	최용국	한국수입자동차협회 이사	융복합 포럼 제도분과 위원
24	윤용원	자동차안전연구원	자율주행자동차 기술 전문가
25	송선영	경상대학교 교수	로봇윤리 전문가
26	신현우	청주교육대학교 교수	로봇윤리 전문가
27	변순용	서울교육대학교 교수	로봇윤리 전문가
28	황창근	홍익대학교 교수	변호사(자율주행관련법)
29	김규현	홍익대학교 교수	융복합 포럼 산업분과 위원
30	손주찬	한국전자통신연구원 본부장	융복합 포럼 생활분과 위원
31	강경표	한국교통연구원 연구위원	융복합 포럼 제도분과 위원
32	장정아	아주대학교 교수	융복합 포럼 제도분과 위원

· 1차 델파이 자문위원 명단

NO	성명	소속	비고
1	윤영한	한국기술교육대학교 교수	융복합 포럼 제도분과 위원장
2	윤일수	아주대학교 교수	융복합 포럼 기술분과 위원장
3	임현준	홍익대학교 교수	융복합 포럼 산업분과 위원장
4	박용훈	교통문화운동본부 대표	융복합 포럼 생활분과 위원
5	변순용	서울교육대학교 교수	로봇윤리 전문가

· 2차 델파이 전문가 의견수렴 명단

NO	성명	소속	비고
1	윤일수	아주대학교 교수	융복합 포럼 기술분과 위원장
2	김범준	LG경제연구원 책임연구원	융복합 포럼 산업분과 위원
3	남두희	한성대학교 교수	융복합 포럼 산업분과 위원
4	박요한	삼성교통안전문화연구소 수석연구원	융복합 포럼 산업분과 위원
5	임현준	홍익대학교 교수	융복합 포럼 산업분과 위원장
6	김규현	홍익대학교 교수	융복합 포럼 산업분과 위원
7	김정룡	한양대학교 교수	융복합 포럼 생활분과 위원장
8	손주찬	한국전자통신연구원 본부장	융복합 포럼 생활분과 위원
9	윤영한	한국기술교육대학교 교수	융복합 포럼 제도분과 위원장
10	이항구	한국산업연구원 선임연구위원	융복합 포럼 제도분과 위원
11	최용국	한국수입자동차협회 이사	융복합 포럼 제도분과 위원
12	이중기	홍익대학교 교수	융복합 포럼 제도분과 위원
13	강경표	한국교통연구원 연구위원	융복합 포럼 제도분과 위원
14	송선영	경상대학교 교수	로봇윤리 전문가
15	황창근	홍익대학교 교수	변호사(자율주행관련법)

· 2차 델파이 자문위원 명단

NO	성명	소속	비고
1	윤영한	한국기술교육대학교 교수	융복합 포럼 제도분과 위원장
2	윤일수	아주대학교 교수	융복합 포럼 기술분과 위원장
3	임현준	홍익대학교 교수	융복합 포럼 산업분과 위원장
4	김정룡	한양대학교 교수	융복합 포럼 기술분과 위원장
5	박용훈	교통문화운동본부 대표	융복합 포럼 생활분과 위원
6	변순용	서울교육대학교 교수	로봇윤리 전문가

부록 3 집필진 및 기관 소개

· 국토교통부 첨단자동차과 : 이창기, 이정규, 이은정

국토교통부 첨단자동차과는 우리나라 자율주행자동차 관련 국내 정책, 법제도 마련 및 자율주행 인프라 구축 업무를 담당하고 있으며, 안전하고 편리한 자율주행 시대를 앞당기고 미래자동차를 선도하기 위해 자율주행자동차 상용화 사업 및 관련 국가 R&D 사업을 총괄하고 있음.

· 홍익대학교 도시공학과 : 황기연, 송재인, 강민희, 조윤지, 이우섭

2017년 04월부터 4년간 국토교통부의 자율주행자동차 국가 R&D 사업을 통해 자율주행자동차의 윤리와 관련된 연구를 수행하고 있으며, 최종적으로 자율주행 시대에 예상되는 문제에 대한 지침으로 윤리 가이드라인을 작성하였음.

· 자동차안전연구원 : 신재곤, 윤용원, 김종화, 조성우, 조광상

자동차안전연구원은 한국교통안전공단의 부설 연구기관으로 1987년 설립되었으며, '안전하며 편리하고 쾌적한 교통환경을 조성하여 모든 국민이 행복한 세상을 만든다'는 미션 아래 제작결함조사, 자동차안전기준국제조화, 자동차안전도평가, 미래자동차연구 등 우리나라 자동차 안전을 책임지고 있으며, 정부 R&D 수행 등 자율주행자동차 상용화 위한 정부의 정책 및 법제도를 지원하고 있음.

· 한국교통연구원 : 김규욱, 문영준, 조선아

미래차교통연구센터는 미래자동차 기반의 교통체계와 정책을 선제적으로 마련하고, 관련 기술과 서비스, 국제표준 등을 개발하고 있으며, 2017년 4월부터 국토교통부의 자율주행자동차 국가 R&D 사업을 통해 자율주행자동차의 사회적 수용성 향상을 위한 기반 기술을 개발함으로써 자율주행차 조기 상용화에 기여하고 있음.