

인공지능 로봇의 위험성 유형별 측정지표 개발

Development of Measurement Indicators by Type of Risk of AI Robots

송 현 경^{1*}

Hyun-kyoung Song

요 약

인공지능 로봇의 산업화가 활발해짐에 따라 윤리적, 기술적 문제가 심각해지고 있음에도 위험성에 관한 연구는 미흡하다. 이러한 실정에서 연구자는 인공지능 로봇의 신체, 권리, 재산, 사회적 위험성을 측정할 수 있는 검증된 지표 52개를 개발하였다. 인공지능 로봇의 위험성 유형별 측정지표 개발을 위해 IRB 심의 이후, 전문가 11명에게 심층 면접하였다. 또한 현장성 검증을 위해 인공지능 로봇이 도입될 수 있는 여러 분야 종사자 328명에게 설문 조사하였으며, 타당성 및 신뢰성 측정을 위해 탐색적 요인분석과 신뢰도 분석, 변수 계산을 위한 상관관계 분석과 다중 회귀분석 등 통계 검증하였다. 이 논문에서 제시하는 측정지표가 표준화된 인공지능 로봇의 개발·인증·교육·정책 등에 널리 활용되어, 사회적으로 공감받고 안전한 인공지능 로봇 산업화의 초석이 되기를 기대한다.

☞ 주제어 : 인공지능 로봇, 인공지능 로봇 위험성 측정 모형, 인공지능 로봇 측정지표, 인공지능 윤리원칙, 로봇 윤리

ABSTRACT

Ethical and technical problems are becoming serious as the industrialization of artificial intelligence robots becomes active, research on risk is insufficient. In this situation, the researcher developed 52 verified indicators that can measure the body, rights, property, and social risk of artificial intelligence robots. In order to develop measurement indicators for each type of risk of artificial intelligence robots, 11 experts were interviewed in-depth after IRB deliberation. In addition, 328 workers in various fields where artificial intelligence robots can be introduced were surveyed to verify their fieldwork, and statistical verification such as exploratory factor analysis, reliability analysis, correlation analysis, and multiple regression analysis was verified to measure validity and reliability. It is expected that the measurement indicators presented in this paper will be widely used in the development, certification, education, and policies of standardized artificial intelligence robots, and become the cornerstone of the industrialization of artificial intelligence robots that are socially sympathetic and safe.

☞ keyword : AI Robot, AI Robot Risk Measurement Model, AI Robot Measurement Indicators, AI Ethical Principles, Robot ethics

1. 서 론

최근 농산물 선별 작업장에서 센서를 점검하고 있던 작업자를 종이 상자도 오인한 로봇이 집게로 압착하면서 인간이 사망하는 사건이 발생했다[1]. 이를 전문가들은 지능형 로봇의 작업으로 인한 기술적 부작용으로 해석한다.

소셜 로봇은 사용자의 개인정보 등을 저장하고 있을 가능성이 높아 폐기할 때 보안 문제가 있다[2]. 인공지능 로봇은 여러 분야에서 다양한 형태로 나타날 수 있으며, 윤리적 문제 외에도 발생할 수 있는 여러 가지 문제들을 현재로서는 모두 예측할 수 없다. 인간 생활을 윤택하게 해주는 인공지능 로봇을 더 안전하게 사용하기 위해서는

윤리적·기술적 문제를 모두 측정할 수 있는 지표 연구가 활발히 이뤄져야 한다. 그동안 미흡했던 국내외의 연구 실정과 향후 연구 방향을 볼때, 인공지능 로봇의 위험성 유형별 측정지표 개발 연구는 매우 중요하다고 할 것이다.

2. 이론적 배경

2.1 인공지능 로봇 윤리 현황

인공지능 로봇 개발이 본격화되면서 기업과 정부에서 인공지능 윤리 원칙이나 인공지능 로봇 개발 안내서를 활발하게 제작하고 있다. 하루가 달리 발전해가는 인공지능 로봇의 산업화로 새로운 위험 요소들이 생겨나고 있다. 인공지능 로봇의 위험 요소는 윤리적인 문제 뿐만 아니라 기술적인 문제들을 포함하고 있다. 국내외 여러 국가는 인공지능 로봇이 해를 끼칠 것에 대비하여 로봇

¹ National Police Agency Seoul Metropolitan Police Agency, Seoul, Korea
* Corresponding author (onlyone_teacher@naver.com)

[Received 6 June 2024, Reviewed 29 June 2024, Accepted 11 July 2024]

관련 법, 인공지능 개발 가이드라인 등을 제시하고 있다. 국내의 인공지능 로봇 윤리 현황은 표 1과 같다.

(표 1) 국내외 인공지능 로봇 윤리 현황
(Table 1) Domestic-Foreign AI Robot Ethics Status

연도	국내외 현황	참고문헌
1942년	아시모프의 로봇 3원칙	
2005년	EU 윤리 로봇 프로젝트	
2007년	국내, 산업자원부 로봇 윤리현장	
2010년	영국, 로봇 원칙	
2014년	EU, 로보틱스 규제 가이드라인	
2016년	일본, 인공지능 R&D 가이드라인 유럽의회, 로봇 윤리 대원칙	
2017년	EU, 로봇법 권고안 국내, 인공지능 윤리 이슈 보고서	(3), (4), (5), (6), (7), (8)
2019년	EU, 신뢰할 수 있는 AI 윤리 가이드라인 중국, AI 윤리 가이드라인	
2020년	미국, 인공지능 사용에 관한 5가지 윤리 지침	
2021년	UNESCO, 인공지능 윤리 권고	
2023년	국내, 과학기술정보통신부 신뢰할 수 있는 인공지능 개발 안내서	

해외 여러 국가들중에서 특히 유럽연합에서 인공지능 로봇과 관련한 정책이 활발하다. 하지만 최근까지 대부분 국가들의 정책을 살펴본 결과 인공지능 로봇의 기술적인 문제를 포함해서 점검할 수 있는 정책보다는 윤리 관련 정책 위주에 초점이 맞춰져 있는 것으로 나타났다.

2.2 국내외 선행연구 고찰

인공지능 로봇 관련 연구 사례를 살펴보기 위해 국내 학술지 검색시스템과 스콜라 구글에서 문헌을 검색하였다.

국내의 인공지능 로봇 관련 선행연구는 표 2와 같이 정리하였다.

(표 2) 국내외 인공지능 로봇 관련 선행연구
(Table 2) Domestic-Foreign AI Robot Prior research

연구 분야	국내외 현황	참고문헌
윤리원칙	· 가정용 헬스 케어 인공지능 로봇의 개발자용 윤리 체크리스트 개발 연구 · 인공지능 로봇 윤리 가이드라인 연구 - 인공지능 로봇 윤리의 4 원칙을 중심으로 등 · A Study on the Future and Ethical Perspectives of Robotics and AI	(4), (9), (10), (11), (12), (13), (14), (15), (16), (17),

연구 분야	국내외 현황	참고문헌
위험성	· The global Landscape of AI ethics guidelines · 배달 로봇의 보도 주행 허용 관련, 위험성 평가 모델을 통해 안전 요구사항 개발 연구 · Toward Safe and Efficient Human - Robot Interaction via Behavior-Driven Danger Signaling	(18)
법, 규범	· 인공지능 로봇에 관한 법적 고찰 · 인공지능 로봇에 관한 형사책임, 형법에 의한 형사 처벌 가능성 · 인공지능과 규범의 형성 · Liability for robots I: legal challenges	
기술활용	· 로봇 산업 분야의 인공지능 활용 현황과 전망 · 로봇과 인간, 로봇과 로봇, 로봇과 장비 및 인간 간의 협업을 위한 인공지능 기술 · Participating in Robots? AI's Role in the Service · The Network Robot System: Enabling Social Human Robot Interaction in Public Spaces	

국내에는 법·규범·기술 활용 방안 연구가 대부분이었고, 개발자용 체크리스트·감정평가 산업·홈 헬스 케어 등의 한정적으로만 적용할 수 있어 범용적인 활용성은 없다는 한계가 있었다. 해외도 인공지능 로봇의 위험성에 관한 연구·성능 향상·인공지능 윤리원칙·기술 활용 등이었다. 하지만 인공지능 로봇의 위험성을 윤리·기술적으로 측정할 수 있는 지표에 관한 연구는 미흡한 실정이었다.

3. 인공지능 로봇의 위험성 유형별 측정지표 개발

3.1 연구방법

인공지능 로봇의 위험성 유형별 측정지표 개발을 위해 성균관대학교 기관 IRB 연구 승인(SKKU 2023-07-026)을 받았다. 2023년 7월 18일부터 7월 28일까지 전문가 심층 면접을 3회기 진행하였다. 2023년 8월 1일부터 8월 15일까지 현장성 검증을 위한 설문조사 결과를 토대로 해서 탐색적 요인분석·신뢰도 분석 및 다중회귀 분석하였다.

3.2 전문가 심층 면접

인공지능 로봇의 위험성 유형별 측정지표 개발을 위해 인공지능 로봇의 위험성 예방을 위한 측정지표 연구에서 도출된 측정지표들을 토대로 확장 연구하였다[24].

당시 델파이 설문조사에 참여했던 전문가 30명 중에서 인공지능 로봇의 위험성 유형별 측정지표 개발을 위한 전문가 심층 면접에 동의한 11명이 참여하였다.

전문가 심층 면접 참여자에 대한 정보는 표 3과 같다.

(표 3) 전문가 심층 면접 참여자 정보
(Table 3) Expert in-depth interview participant information

구분		전문 분야				인원수	
		IT, SW	AI, 윤리	공공	로봇	명	%
성별	남	3	1	4	2	10	90.9
	여	-	1	-	-	1	9.1
	소계	3	2	4	2	11	100
연령	30대	2	-	2	-	4	36.4
	40대	-	1	-	1	2	18.2
	50대	1	1	2	1	5	45.5
	소계	3	2	4	2	11	100
직업	개발·기획	2	-	4	1	7	63.6
	연구원	-	1	-	1	2	18.2
	교수	1	1	-	-	2	18.2
	소계	3	2	4	2	11	100
직업 및 연구 경력	5년~10년	2	-	2	-	4	36.4
	11년~20년	-	1	1	1	3	27.3
	21년~30년	1	-	1	1	3	27.3
	31년 이상	-	1	-	-	1	9.1
	소계	3	2	4	2	11	100

기존 인공지능 로봇의 위험성 예방을 위한 측정지표 연구에서는 4개 유형 및 11개 항목별 37개 측정지표가 도출되었다. 이 연구에 참여한 전문가들을 심층 면접한 결과 총 65개 측정지표를 도출하였다. 전문가 심층 면접 3회차를 진행하여 현장성 검증을 위한 설문 방법을 선정 하고, 연구가설 검증을 위한 설문지 문항을 설계하였다.

3.3 변수명 정리 및 연구가설 설정

인공지능 로봇의 위험성 유형별 항목 및 측정지표의 연구가설 설정을 위한 변수명은 표 4와 같이 정리하였다.

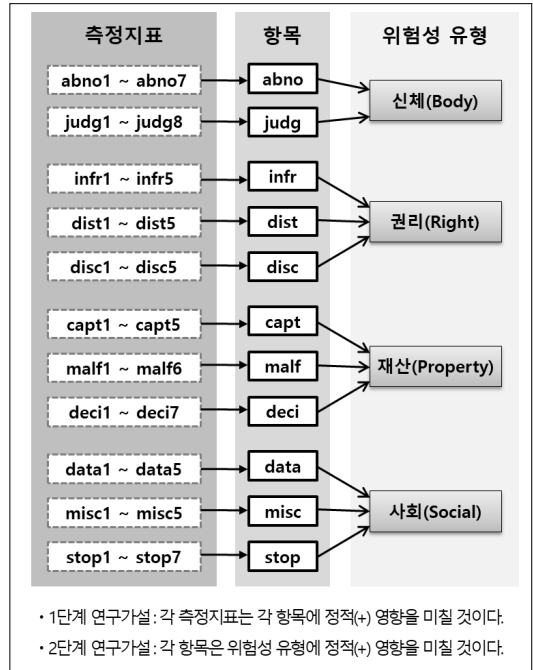
(표 4) 유형별 항목 및 측정지표 변수명 정리
(Table 4) Organize item and metric variable names by type

유형	항목	측정지표	변수명
인간을 해칠 판단 (Body)	운영, 작업 중 작동 이상 (abno)	센서 및 인식 오류	abno1
		운동 지능 제어 실패	abno2
		기체 성능 저하	abno3
		프로그램 오류로 인한 오작동	abno4
		통신 장애 등 사고 위험 감지 지연	abno5
		모션 제어 시스템 성능 저하	abno6
		부품 교체 시기 경과	abno7
신체에 대한 물리적 위험성 (Body)	인간을 해칠 판단 (judg)	의도 파악 인식 오류	judg1
		협동 작업 중, 규칙 판단 지능의 오류	judg2
		프로그램의 판단 오류로 사고 희생자 선택 착오	judg3
		학습량 부족으로 맞지 않는 의학적 자료 제공	judg4
		학습된 규칙이나 로봇의 원칙을 넘어서는 판단	judg5
		라벨링 오류로 인한 잘못된 정보 제공	judg6
		장애물 인식 오류	judg7
		비상 상황 안전 정지 모니터링 실패	judg8
인간의 권리에 대한 위험성 (Right)	인격권 침해 (infr)	명령어 등 음성을 잘못 인식	infr1
		감정 피드백 오류	infr2
		대상 인식 판단 오류	infr3
		도덕성 검증이 되지 않은 학습 데이터 사용	infr4
		멀티 모달리티 오류	infr5
자유 의사 방해 (dist)	자유의사 방해 (dist)	의사 결정을 특정 방향으로 유도	dist1
		어른 형성 조장	dist2
		중요한 결정을 위한 관리자의 개입 차단	dist3
		필터링 오류로 의사표현의 자율성 침해	dist4
		우월적인 지위를 이용해서 인간의 심리 지배	dist5
선택적 서비스, 정보 제공 (disc)	선택적 서비스, 정보 제공 (disc)	학습 데이터 오류로 인해 잘못된 언어 사용	disc1
		편향된 학습 데이터로 인한 차별적 정보 제공	disc2
		시스템 공격 등으로 개인정보의 침해나 유출	disc3
		대상자의 지식수준에 따라 선별적 정보 제공	disc4
		시나리오 설계 문제로 부적절한 콘텐츠 제공	disc5

유형	항목	측정지표	변수명
재산 손실의 위험성 (Property)	정보 입력, 탈취 (capt)	시스템 접근 제어망 침범	capt1
		비인가 시스템 접근 허용	capt2
		악성코드, 바이러스 감염 등으로 시스템 오류	capt3
		데이터 손실 방지 시스템 취약	capt4
		허용되지 않은 정보 수집 및 이용	capt5
	기계 결함 오작동 (malf)	센서 모니터링 실패	malf1
		다양한 상황에 의한 충돌로 시설물 파손	malf2
		모터, 유압 등 구동부 문제	malf3
		정기검사 미 점검 등 정비 불량	malf4
		감속기 제어의 문제	malf5
		제어기 오류	malf6
	잘못된 의사 결정 (deci)	유형의 재물 훼손	deci1
		인식 지능에 대한 오류	deci2
		창작물 무단 복제	deci3
		하자 상품 추천	deci4
		안전거리 계산 오류	deci5
		알고리즘의 최적화 성능 부족	deci6
		사물 인식 판단 능력 저하	deci7
	사회적 위험성 (Social)	데이터 문제 (data)	데이터 유출
개인 등 정보 무작위 수집			data2
부적절한 학습 데이터 수집			data3
습득한 정보를 권한 없는 자에게 제공			data4
허위 정보 제공			data5
법규 판단 오류 (misc)		법률, 규칙 혼동	misc1
		인간이 의도하지 않은 행위로 혼란유발	misc2
		차선, 신호 등 인식 오류	misc3
		법규 최신화 업데이트 오류	misc4
		허용되지 않은 권한의 오인 판단	misc5
도로상 작동 중지 (stop)		관리 부실 등 연료 부족으로 정상 운행 중 작동 정지	stop1
		통신 이상 등 장애	stop2
		보행자 통행 방해	stop3
		도로매칭 알고리즘 오류	stop4
		주변 환경 변화에 대한 반응 속도저하	stop5
		이동 속도 제어나 조절 실패로 운행 정지	stop6
		데이터 업데이트 오류	stop7

인공지능 로봇의 위험성 유형별 측정지표 개발을 위해 1단계로 측정지표들이 각 항목에 영향을 미치는지 검증하기 위한 연구가설을 설정하였다. 2단계에서는 각 항목이 위험성 유형에 영향을 미치는지 검증하기 위한 연구가설을

설정하였다. 인공지능 로봇의 위험성 유형별 항목 및 측정지표의 연구가설은 다음의 그림 1과 같이 정리하였다.



(그림 1) 연구가설

(Figure 1) Research hypothesis

4. 설문조사 통계 분석 및 연구 결과

연구가설 검증을 위해 전문가 심층 면담에서 독립변수-종속변수의 인과관계를 확인하는 다중 회귀분석 방법이 선정되었다. 탐색적 요인분석 및 신뢰도 분석으로 통계 검증하였고, 그 결과를 토대로 상관관계 분석 및 변수를 다시 계산하여 다중 회귀 분석하였다.

4.1 인구 통계학적 특성

현장성 검증을 위해 인공지능 로봇이 도입될 수 있는 여러 분야 종사자 339명에게 리커트 5점 척도를 활용한 설문 조사를 진행하였다. 개인 정보 수집 등에 동의하지 않은 11명의 조사지는 제외하고, 유효수 328명이 응답한 설문조사 결과로 통계 분석하였다. 현장성 검증을 위한 설문조사 참여자의 인구통계학적 특성은 표 5와 같다.

(표 5) 현장성 검증 설문 참여자의 인구통계학적 특성
(Table 5) Demographic characteristics of field verification survey participants

구분		빈도수(명)	구성비율(%)	누적 비율(%)
성별	남	226	68.90	68.90
	여	102	31.10	100
	소계	328	100	100
연령	20 ~ 29세	28	8.54	8.54
	30 ~ 39세	80	24.39	32.93
	40 ~ 49세	133	40.55	73.48
	50 ~ 59세	65	19.82	93.30
	60세 이상	22	6.71	100
	소계	328	100	100
직업 분류	공무원·교사 등 공직	103	31.40	31.40
	ICT분야 연구직 등	97	29.57	60.97
	제조·기술·생산직	58	17.68	78.65
	자영업·서비스직	54	16.46	95.11
	농·어업	16	4.89	100
	소계	328	100	100
경력	1 ~ 5년	53	16.16	16.16
	5 ~ 10년	63	19.21	35.37
	10 ~ 15년	50	15.24	50.61
	15 ~ 20년	61	18.60	69.21
	20 ~ 25년	60	18.29	87.50
	25 ~ 30년	25	7.62	95.12
	30년 이상	16	4.88	100
	소계	328	100	100

4.2 탐색적 요인분석 및 신뢰도 분석결과

전문가 심층 면접 등 연구를 통해 도출된 측정지표의 탐색적 요인분석 적합성을 확인하고자 각 유형별로 표준형성의 적절성을 나타내는 KMO 및 Bartlett의 구형성을 검증하였다.

신체에 대한 물리적 위험성 유형의 검증 결과는 표 6과 같이 KMO는 0.944로 나타났고, Bartlett의 구형성 검증의 유의확률은 0.000으로 탐색적 요인분석에 문제가 없었다.

(표 6) 신체에 대한 물리적 위험성 유형의 KMO 및 Bartlett의 검증 결과

(Table 6) Validation results from KMO and Bartlett of the type of physical risk to the body

표준형성 적절성의 Kaiser-Meyer-Olkin 측도		.944
Bartlett의 구형성 검증	근사 카이제곱	3867.978
	자유도	105
	유의확률	0.000

신체에 대한 물리적 위험성 유형의 탐색적 요인분석 및 신뢰도 분석결과는 표 7과 같다.

(표 7) 신체에 대한 물리적 위험성 유형의 탐색적 요인분석 및 신뢰도 분석결과

(Table 7) Exploratory factor analysis and reliability analysis of physical risk types for the body

변수 명	요인 적재량		공동성	고유 값	누적률	신뢰도	
	judg	abno				D*	**
judg5	.839	.235	.759	5.564	37.096	.931	.940
judg6	.799	.254	.703			.933	
judg4	.788	.235	.677			.935	
judg2	.776	.404	.766			.929	
judg8	.773	.346	.717			.932	
judg7	.771	.339	.710			.932	
judg1	.765	.325	.691			.933	
judg3	.723	.367	.657			.935	
abno5	.269	.830	.643	4.879	69.624	.901	.920
abno4	.239	.811	.714			.907	
abno2	.303	.803	.737			.904	
abno1	.268	.787	.691			.908	
abno6	.319	.736	.643			.909	
abno7	.345	.720	.637			.909	
abno3	.430	.629	.581			.914	

* D. : Alpha if Item Deleted, ** : Cronbach

신체에 대한 물리적 위험성 유형 15개 변수의 탐색적 요인분석 결과, 2개 항목으로 묶이는 것을 알 수 있었다.

2개 항목은 전문가 심층 면접을 통해 도출된 인간을 해칠 판단의 judg 및 운행 작업 중 작동 이상의 abno로 정의하였다. 인간의 권리에 대한 위험성 유형의 KMO 및 Bartlett 검증 결과는 표 8과 같으며, 표준형성의 적절성 측도 및 구형성 수치가 탐색적 요인분석에 문제없었다.

(표 8) 인간의 권리에 대한 위험성 유형의 KMO 및 Bartlett의 검증 결과

(Table 8) Validation results from KMO and Bartlett of risk types for human rights

표준형성 적절성의 Kaiser-Meyer-Olkin 측도		.949
Bartlett의 구형성 검증	근사 카이제곱	3867.424
	자유도	105
	유의확률	0.000

인간 권리에 대한 위험성 유형의 탐색적 요인분석 및 신뢰도 분석결과는 다음 표 9와 같다. 인간 권리에 대한

위험성 유형의 탐색적 요인분석 및 신뢰도 분석결과 15개 변수는 3개 항목으로 묶였다. 이를 전문가 심층 면접에서 도출된 인격권 침해의 *infr*, 자유의사 방해의 *disc*, 선택적 서비스, 정보 제공의 *disc*로 구분하였다.

(표 9) 인간의 권리에 대한 위험성 유형의 탐색적 요인분석 및 신뢰도 분석결과

(Table 9) Exploratory factor analysis and reliability analysis of risk types for human rights

변수 명	요인 적재량			공통성	고유 값	누적률	신뢰도	
	<i>infr</i>	<i>disc</i>	<i>disc</i>				D *	**
<i>dist5</i>	.762	.310	.251	.739	3.742	24.949	.895	.910
<i>dist2</i>	.741	.301	.337	.754			.887	
<i>dist4</i>	.741	.358	.347	.797			.878	
<i>dist1</i>	.711	.265	.372	.714			.894	
<i>dist3</i>	.642	.331	.396	.678			.895	
<i>infr3</i>	.244	.838	.241	.819	3.732	49.831	.896	.919
<i>infr2</i>	.295	.817	.253	.818			.895	
<i>infr1</i>	.280	.808	.252	.795			.898	
<i>infr5</i>	.444	.672	.302	.739			.899	
<i>infr4</i>	.455	.563	.376	.666			.912	
<i>disc2</i>	.327	.208	.777	.755	3.565	73.598	.856	.888
<i>disc3</i>	.260	.171	.735	.636			.884	
<i>disc5</i>	.293	.380	.707	.730			.856	
<i>disc1</i>	.270	.415	.676	.702			.862	
<i>disc4</i>	.428	.243	.674	.697			.859	

* D. : Alpha if Item Deleted, ** : Cronbach

재산 손실의 위험성 유형의 검증 결과는 다음 표 10과 같으며, 표준형성의 적절성 측도 및 구형성 검증 수치는 탐색적 요인분석에 문제없는 것으로 나타났다.

(표 10) 재산 손실의 위험성 유형의 KMO 및 Bartlett의 검증 결과

(Table 10) Validation results from KMO and Bartlett of the types of risk of property loss

표준형성 적절성의 Kaiser-Meyer-Olkin 측도		.952
Bartlett의 구형성 검증	근사 카이제곱	4976.460
	자유도	153
	유의확률	0.000

재산 손실 위험성 유형의 탐색적 요인분석 및 신뢰도 분석결과는 다음 표 11과 같다. 재산 손실 위험성 유형 18개 변수의 탐색적 요인분석 결과, 3개 항목으로 묶였다.

전문가 심층 면접에서 도출된 잘못된 의사결정의 *deci*, 기계 결함 오작동의 *malf*, 정보 입력, 탈취인 *capt*로 구분하였다.

(표 11) 재산 손실 위험성 유형의 탐색적 요인분석 및 신뢰도 분석결과

(Table 11) Exploratory Factor Analysis and Reliability Analysis of Property Loss Risk Types

변수 명	요인 적재량			공통성	고유 값	누적률	신뢰도	
	<i>deci</i>	<i>malf</i>	<i>capt</i>				D *	**
<i>deci3</i>	.805	.257	.117	.728	4.783	26.571	.925	.932
<i>deci4</i>	.782	.324	.189	.752			.921	
<i>deci2</i>	.725	.320	.370	.765			.919	
<i>deci1</i>	.723	.255	.334	.700			.924	
<i>deci6</i>	.717	.282	.326	.699			.923	
<i>deci7</i>	.708	.289	.361	.715	.922	4.334	50.649	.924
<i>deci5</i>	.678	.376	.336	.713	.921			
<i>malf4</i>	.317	.780	.289	.793	.905			
<i>malf3</i>	.308	.749	.268	.727	.911			
<i>malf2</i>	.329	.738	.290	.737	.910			
<i>malf5</i>	.327	.728	.324	.742	.909	4.116	73.513	.920
<i>malf6</i>	.334	.689	.378	.729	.910			
<i>malf1</i>	.280	.680	.320	.644	.919			
<i>capt1</i>	.271	.289	.780	.766	.900			
<i>capt4</i>	.274	.363	.767	.796	.897			
<i>capt5</i>	.285	.303	.763	.756	.904	4.116	73.513	.920
<i>capt3</i>	.285	.264	.749	.712	.911			
<i>capt2</i>	.298	.364	.733	.759	.901			

* D. : Alpha if Item Deleted, ** : Cronbach

사회적 위험성 유형의 검증 결과는 표 12와 같으며, 표준형성의 적절성 측도 및 구형성 검증 수치는 탐색적 요인분석에 문제없는 것으로 나타났다.

(표 12) 사회적 위험성 유형의 KMO 및 Bartlett의 검증 결과 (Table 12) Results of KMO and Bartlett's verification of social risk types

표준형성 적절성의 Kaiser-Meyer-Olkin 측도		.947
Bartlett의 구형성 검증	근사 카이제곱	5431.067
	자유도	136
	유의확률	0.000

사회적 위험성 유형의 탐색적 요인분석 및 신뢰도 분석 결과는 다음 표 13과 같다.

(표 13) 사회적 위험성 유형의 탐색적 요인분석 및 신뢰도 분석결과

(Table 13) Exploratory factor analysis and reliability analysis of social risk types

변수 명	요인 적재량			공통성	고유 값	누적률	신뢰도	
	stop	data	misc				D*	**
stop4	.814	.283	.229	.795	5.366	31.565	.940	.949
stop2	.809	.300	.271	.818			.939	
stop5	.796	.271	.300	.797			.940	
stop1	.786	.331	.258	.793			.941	
stop7	.768	.294	.268	.748			.943	
stop3	.743	.255	.310	.714			.945	
stop6	.743	.299	.289	.724	.944	54.679	.927	
misc5	.335	.786	.309	.826	.904			
misc2	.328	.763	.373	.829	.901			
misc1	.328	.752	.307	.766	.916			
misc4	.373	.730	.294	.758	.913			
misc3	.322	.662	.403	.704	.919			
data2	.306	.336	.800	.847	3.791	76.978	.887	.919
data3	.263	.273	.790	.768			.908	
data4	.305	.299	.781	.793			.896	
data5	.411	.382	.635	.719			.904	
data1	.330	.420	.634	.687			.907	

* D. : Alpha if Item Deleted, ** : Cronbach

사회적 위험성 유형의 17개 변수를 탐색적 요인분석한 결과, 3개 항목으로 묶였다. 이를 전문가 심층 면접을 통해서 도출된 도로상 작동 중지의 stop, 데이터 문제의 data, 법규 판단 오류의 misc로 구분하였다.

4.3 상관관계 분석

연구 가설 검정을 위해 변수계산하였다. 탐색적 요인 분석, 신뢰도 분석 등 변수정제 과정을 통해 최종적으로 남아 있는 변수들을 모두 더해 변수의 갯수만큼 나누고 하나의 변수로 만들어주는 상관관계 분석을 하였다.

신체에 대한 물리적 위험성 유형의 요인별 상관관계를 분석한 결과, abno와 judg의 상관관계는 0.691의 유의미한 수치로 나타났으며, 다음의 표 14와 같다.

(표 14) 신체에 대한 물리적 위험성 유형의 상관분석 결과
(Table 14) Correlation analysis of the types of physical risks to the body

변수명	평균	표준편차	상관관계	
			abno	judg
abno	4.2282	.60634	1	
judg	4.0998	.71992	.691**	1

** 상관계수는 0.01 수준(양쪽)에서 유의합니다.

인간의 권리에 대한 위험성 유형의 요인별 상관관계 분석결과는 표 15와 같다.

infr와 dist는 0.763의 상관관계가 있었고, infr와 disc의 상관관계는 0.717로 나타났다. dist와 disc의 상관관계는 0.781로 모든 요인의 상관관계가 유의미하게 나타났다.

(표 15) 인간의 권리에 대한 위험성 유형의 상관분석 결과
(Table 15) Correlation of Risk Types on Human Rights Results

변수명	평균	표준편차	상관관계		
			infr	dist	disc
infr	4.0232	.77370	1		
dist	4.1567	.71005	.763**	1	
disc	4.2018	.63022	.717**	.781**	1

** 상관계수는 0.01 수준(양쪽)에서 유의합니다.

재산 손실의 위험성 유형의 요인별 상관관계를 분석한 결과는 표 16과 같다.

capt와 malf의 상관관계는 0.747로 나타났고, capt와 deci은 0.707의 유의미한 수치로 나타났다. malf과 deci의 상관관계는 0.750로 모든 요인의 상관관계가 유의미하게 나타났다.

(표 16) 재산 손실 위험성 유형의 상관분석 결과
(Table 16) Correlation Analysis Results of Property Loss Risk Types

변수명	평균	표준편차	상관관계		
			capt	malf	deci
capt	4.2640	.66649	1		
malf	4.2571	.64381	.747**	1	
deci	4.1655	.70198	.707**	.750**	1

** 상관계수는 0.01 수준(양쪽)에서 유의합니다.

사회적 위험성 유형의 요인별 상관관계 분석결과는 다음의 표 17과 같다.

(표 17) 사회적 위험성 유형의 상관분석 결과
(Table 17) Correlation Analysis Results of Types of Social Risk

변수명	평균	표준편차	상관관계		
			data	misc	stop
data	4.1628	.71640	1		
misc	4.1506	.74218	.788**	1	
stop	4.3184	.69385	.721**	.740**	1

** 상관계수는 0.01 수준(양쪽)에서 유의합니다.

data과 misc의 상관관계는 0.788의 수치로 나타났으며, data과 stop 상관관계는 0.721로 나타났다. misc와 stop의 상관관계는 0.74로 모든 요인의 상관관계가 유의미하게 나타났다.

4.4 다중회귀 분석

여러 개 독립변수가 한 개 종속변수에 미치는 영향을 확인하고자 다중 회귀분석을 실시하였다. 신체에 대한 물리적 위험성 유형의 다중회귀 분석결과는 표 18과 같다.

(표 18) 신체에 대한 물리적 위험성 유형의 다중회귀 분석결과
(Table 18) Multiple-term analysis of the type of physical risk to the body

종속변수	독립변수	표준 오차	베타 ()값	t값	유의 확률	공차 한계
abno	상수	.027		1.872	.062	
	abno1	.008	.241	24.004	.000**	.405
	abno2	.009	.154	14.184	.000**	.347
	abno3	.008	.223	23.954	.000**	.471
	abno4	.008	.210	20.177	.000**	.378
	abno5	.009	.175	15.756	.000**	.330
	abno6	.008	.199	19.799	.000**	.406
	R = .993α, R ² = .987, 수정된 R ² = .987 F = 4025.847, p = .000**, Durbin-Watson = 1.833					
judg	상수	.032		1.406	.161	
	judg1	.011	.211	17.130	.000**	.374
	judg2	.012	.200	14.575	.000**	.302
	judg3	.010	.169	14.530	.000**	.420
	judg4	.009	.219	19.534	.000**	.453
	judg6	.010	.208	17.187	.000**	.390
	judg7	.011	.167	13.301	.000**	.363
	R = .991α, R ² = .982, 수정된 R ² = .981 F = 2872.916, p = .000**, Durbin-Watson = 1.912					

은행 작업 중, 작동 이상의 abno 변수 간 및 인간을 해칠 판단의 judg 변수 간 다중 회귀분석 결과, 수정된 R²이 0.987, 0.981로 회귀선의 설명력은 좋은 것으로 나타났다.

Durbin-Watson 값도 2에 가까워 잔차의 독립성도 확립되었다. 종속변수의 분산분석 F 통계량 유의확률 p 값이 0.000으로 회귀식은 유의하다는 연구가설이 채택되었다.

인간 권리에 대한 위험성 유형의 다중회귀 분석결과는 표 19와 같다.

(표 19) 인간의 권리에 대한 위험성 유형의 다중회귀 분석결과
(Table 19) Results of multi-session analysis of risk types for human rights

종속변수	독립변수	표준 오차	베타 ()값	t값	유의 확률	공차 한계
infr	상수	.031		2.476	.014	
	infr1	.011	.278	22.587	.000**	.355
	infr2	.010	.345	27.587	.000**	.345
	infr4	.010	.267	24.152	.000**	.442
	infr5	.011	.245	19.716	.000**	.350
	R = .991α, R ² = .983, 수정된 R ² = .982 F = 4561.581, p = .000**, Durbin-Watson = 1.841					
dist	상수	.038		1.812	.071	
	dist1	.012	.283	20.635	.000**	.433
	dist2	.013	.262	17.942	.000**	.383
	dist3	.012	.280	20.272	.000**	.427
	dist4	.014	.304	19.461	.000**	.334
	R = .987α, R ² = .974, 수정된 R ² = .973 F = 2980.427, p = .000**, Durbin-Watson = 1.908					
disc	상수	.038		1.968	.050	
	disc1	.010	.326	24.997	.000**	.467
	disc2	.012	.245	17.526	.000**	.405
	disc3	.009	.301	26.250	.000**	.604
	disc5	.012	.305	22.654	.000**	.438
	R = .987α, R ² = .974, 수정된 R ² = .974 F = 3070.575, p = .000**, Durbin-Watson = 1.967					

*.p < 0.05 **.p < 0.01

인격권 침해의 infr 변수 간, 자유의사 방해의 dist 변수 간, 선택적 서비스나 정보 제공의 disc 변수 간 다중 회귀 분석 결과, 수정된 R²이 0.982, 0.973, 0.974로 회귀선의 설명력은 좋은 것으로 나타났다.

Durbin-Watson 값도 2에 가까워 잔차의 독립성도 확립되었다. 종속변수의 분산분석 F 통계량 유의확률 p 값이 0.000으로 회귀식은 유의하다는 연구가설이 채택되었다.

재산 손실 위험성 유형의 다중회귀 분석결과는 다음의 표 20과 같다.

(표 20) 재산 손실 위험성 유형의 다중회귀 분석결과
(Table 20) Results of multi-session analysis of property loss risk types

종속변수	독립변수	표준 오차	베타 ()값	t값	유의 확률	공차 한계
capt	상수	.035		1.434	.153	
	capt1	.012	.276	20.034	.000**	.342
	capt2	.012	.294	22.440	.000**	.377
	capt3	.011	.256	20.784	.000**	.427
	capt4	.012	.296	21.254	.000**	.334
	R = .989α, R ² = .979, 수정된 R ² = .979 F = 3778.046, p = .000**, Durbin-Watson = 1.962					

종속변수	독립변수	표준 오차	베타 ()값	t값	유의 확률	공차 한계
malf	상수	.030		.218	.828	
	malf1	.008	.233	22.892	.000**	.451
	malf2	.010	.231	21.028	.000**	.387
	malf3	.010	.209	17.979	.000**	.345
	malf4	.010	.257	21.694	.000**	.333
	malf6	.010	.227	20.418	.000**	.379
R = .992 α , R ² = .985, 수정된 R ² = .985 F = 4209.241, p = .000**, Durbin-Watson = 1.929						
deci	상수	.024		2.880	.004	
	deci1	.007	.201	22.513	.000**	.384
	deci2	.009	.198	20.325	.000**	.321
	deci3	.007	.188	21.823	.000**	.411
	deci4	.008	.200	21.615	.000**	.358
	deci5	.008	.184	19.497	.000**	.344
deci6	.008	.201	22.816	.000**	.393	
R = .995 α , R ² = .990, 수정된 R ² = .990 F = 5402.944, p = .000**, Durbin-Watson = 1.970						

*.p < 0.05 **p < 0.01

정보 입력·탈취의 malf, 기계 결함 오작동의 malf, 잘못된 의사 결정의 deci 변수 간 다중 회귀분석 결과, 수정된 R²이 0.979, 0.985, 0.990으로 회귀선의 설명력은 좋은 것으로 나타났다. Durbin-Watson 값도 2에 가까운 잔차의 독립성도 확립되었다.

종속변수의 분산분석 F 통계량 유의확률 p 값이 0.000으로 회귀식은 유의하다는 연구가설이 채택되었다. 사회적 위험성 유형의 다중회귀 분석결과는 표 21과 같다.

(표 21) 사회적 위험성 유형의 다중회귀 분석결과

(Table 21) Results of multi-term analysis of social risk types

종속변수	독립변수	표준 오차	베타 ()값	t값	유의 확률	공차 한계
data	(상수)	.033		.004	.997	
	data1	.011	.270	22.376	.000**	.406
	data2	.014	.284	19.138	.000**	.268
	data3	.010	.295	24.513	.000**	.410
	data5	.010	.284	23.899	.000**	.420
R = .990 α , R ² = .981, 수정된 R ² = .981 F = 4133.713, p = .000**, Durbin-Watson = 1.948						
misc	(상수)	.027		2.262	.024	
	misc1	.009	.338	31.961	.000**	.361
	misc2	.012	.219	16.559	.000**	.231
	misc3	.009	.296	27.637	.000**	.353
	misc4	.009	.277	26.996	.000**	.384
R = .993 α , R ² = .987, 수정된 R ² = .987 F = 6092.839, p = .000**, Durbin-Watson = 1.900						

종속변수	독립변수	표준 오차	베타 ()값	t값	유의 확률	공차 한계
stop	(상수)	.033		3.083	.002	
	stop1	.012	.220	14.687	.000**	.248
	stop2	.013	.262	18.405	.000**	.276
	stop3	.010	.250	21.000	.000**	.396
	stop4	.013	.181	12.235	.000**	.255
	stop7	.012	.205	15.687	.000**	.328
	R = .991, R ² = .982, 수정된 R ² = .982 F = 3509.446, p = .000**, Durbin-Watson = 1.907					

*.p < 0.05 **p < 0.01

데이터 문제의 data, 범규 판단 오류의 misc, 도로상 작동 중지 stop 변수 간 다중 회귀분석 결과, 수정된 R²이 0.981, 0.987, 0.982로 회귀선의 설명력은 좋은 것으로 나타났다. Durbin-Watson 값도 2에 가까운 수치로 나타나 잔차의 독립성도 확립되었다. 종속변수의 분산분석 F 통계량 유의확률 p 값이 0.000으로 회귀식은 유의하다는 연구가설이 채택되었다. 연구 가설 2단계 검증에 위한 위험성 유형별 항목 간 다중회귀 분석결과는 표 22와 같다.

(표 22) 위험성 유형별 항목 간 다중회귀 분석결과

(Table 22) Results of multi-session analysis between items by risk type

종속변수	독립변수	표준 오차	베타 ()값	t값	유의 확률	공차 한계
Body	(상수)	.024		.165	.869	
	abno	.008	.507	66.367	.000**	.522
	judg	.007	.574	75.171	.000**	.522
R = .995 α , R ² = .990, 수정된 R ² = .990 F = 16231.850, p = .000**, Durbin-Watson = 1.973						
Right	(상수)	.023		-.170	.865	
	infr	.007	.391	47.095	.000**	.381
	dist	.008	.354	38.174	.000**	.305
disc	.009	.345	40.170	.000**	.354	
R = .996 α , R ² = .992, 수정된 R ² = .991 F = 12629.894, p = .000**, Durbin-Watson = 1.928						
Pro perty	(상수)	.018		1.434	.152	
	capt	.006	.281	43.365	.000**	.392
	malf	.007	.352	50.893	.000**	.344
deci	.006	.464	71.466	.000**	.389	
R = .997 α , R ² = .995, 수정된 R ² = .995 F = 20183.488, p = .000**, Durbin-Watson = 1.911						
Social	(상수)	.021		-.392	.695	
	data	.008	.324	38.821	.000**	.337
	misc	.008	.351	40.940	.000**	.318
	stop	.007	.416	54.559	.000**	.402
R = .996 α , R ² = .992, 수정된 R ² = .992 F = 14111.029, p = .000**, Durbin-Watson = 1.903						

*.p < 0.05 **p < 0.01

신체에 대한 물리적 위험성의 Body, 인간의 권리에 대한 위험성의 Right, 재산 손실의 위험성의 Property, 사회적 위험성의 Social 변수 간 다중 회귀분석 결과, 수정된 R² 이 0.990, 0.991, 0.995, 0.992로 회귀선의 설명력은 좋은 것으로 나타났다. Durbin-Watson 값도 2에 가까운 수치로 나타나 잔차의 독립성도 확립되었다. 종속변수 분산분석 F 통계량 유의확률 p 값이 0.000으로 나타나 회귀식은 유의하다는 연구가설이 채택되었다.

1단계 연구 가설을 검증한 결과, abno1~abno6은 채택되었고, abno7은 기각되었다. judg1~judg4, judg6~judg7은 채택되었고, judg5, judg8은 기각되었다. infr1~infr2, infr4~infr5는 채택되었고, infr3은 기각되었다. dist1~dist4는 채택되었고, dist5는 기각되었다. disc1~disc3, disc5는 채택되었고, disc4는 기각되었다. capt1~capt4는 채택되었고, capt5는 기각되었다. malf1~malf4, malf6은 채택되었고, malf5는 기각되었다. deci1~deci6은 채택되었고, deci7은 기각되었다. data1~data3, data5는 채택되었고, data4는 기각되었다. misc1~misc4는 채택되었고, misc5는 기각되었다.

stop1~stop4, stop7은 채택되었고, stop5~stop6은 기각됨에 따라 총 65개 측정지표 중에서 52개의 연구가설이 채택되었다. 2단계 연구 가설을 검증한 결과 abno, judg, infr, dist, disc, capt, malf, deci, data, misc, stop 11개의 항목 모두 연구가설이 채택되었다.

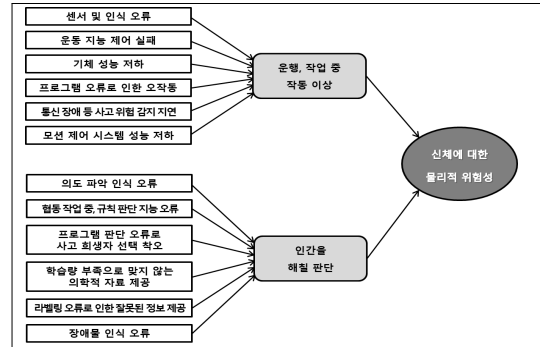
4.5 연구 결과

인공지능 로봇의 위험성 유형별 측정지표 개발을 위해 전문가 심층 면접 및 현장성 설문조사를 하였다. 그 결과를 탐색적 요인분석 및 신뢰도 분석, 상관관계 분석 및 다중 회귀 분석하였다. 연구가설을 통계 검증하여 개발한 4개 유형 및 11개 항목별 52개 측정지표는 그림 2 ~ 5와 같다.

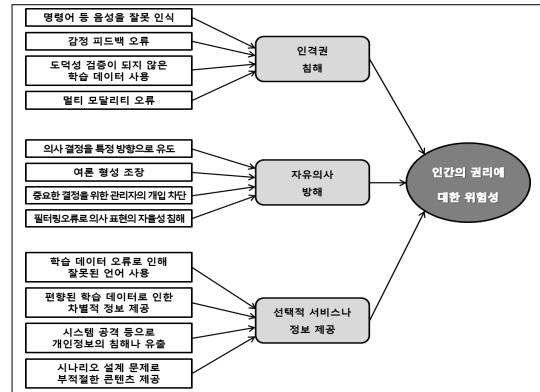
신체에 대한 물리적 위험성은 인공지능 로봇의 운행, 작업 중 작동 이상 및 인간을 해칠 판단에 의해 발생할 수 있다.

따라서 운행, 작업 중에 작동 이상을 일으킬 수 있는 6개 측정지표 및 인간을 해칠 판단을 할 수 있는 6개 측정지표를 활용하여 인공지능 로봇을 점검할 필요가 있음을 그림 2와 같이 제안한다.

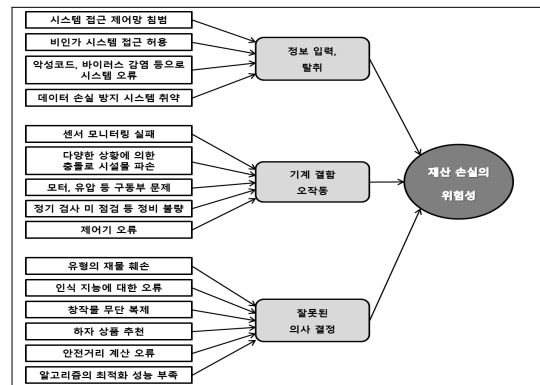
인간 권리에 위험성은 인공지능 로봇의 인격권 침해 및 자유의사 방해, 선택적 서비스나 정보 제공에 의해 발생할 수 있다. 따라서 인격권을 침해할 수 있는 4개 측정지표 및 자유의사를 방해할 수 있는 4개 측정지표 및 선택적 서비스나 정보 제공을 할 수 있는 4개 측정지표를 활용하여 인공지능 로봇을 점검할 필요가 있음을 그림 3과 같이 제안한다.



(그림 2) 신체에 대한 물리적 위험성 유형의 측정지표 모형 (Figure 2) A metric model of the type of physical risk to the body

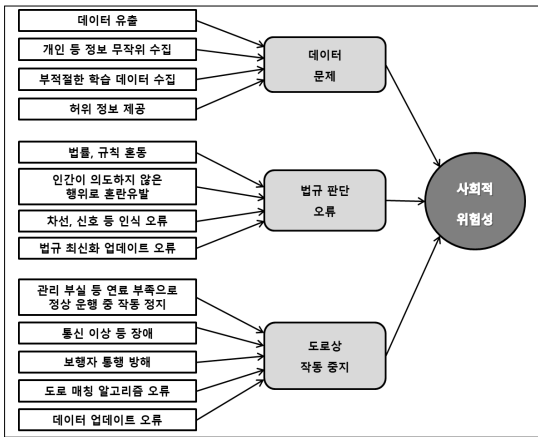


(그림 3) 인간의 권리에 대한 위험성 유형의 측정지표 모형 (Figure 3) A metric model of the type of risk for human rights



(그림 4) 재산 손실 위험성 유형의 측정지표 모형 (Figure 4) Measurement Indicator Model of Property Loss Risk Types

재산 손실의 위험성은 인공지능 로봇의 정보 입력, 탈취 및 기계 결합 오작동, 잘못된 의사 결정에 의해 발생할 수 있다. 따라서 정보 입력이나 탈취를 할 수 있는 4개 측정지표 및 기계 결합으로 인한 오작동을 일으킬 수 있는 5개 측정지표, 잘못된 의사 결정을 할 수 있는 6개 측정지표를 활용하여 인공지능 로봇을 점검할 필요가 있음을 그림 4와 같이 제안한다. 사회적 위험성은 인공지능 로봇의 데이터 문제 및 법규 판단 오류, 도로상 작동 중지에 의해 발생할 수 있다. 따라서 데이터 문제를 일으킬 수 있는 4개 측정지표, 법규 판단에 오류를 일으킬 수 있는 4개 측정지표, 도로상에서 작동이 중지될 수 있는 5개 측정지표를 활용하여 인공지능 로봇을 점검할 필요가 있음을 그림 5와 같이 제안한다.



(그림 5) 사회적 위험성 유형의 측정지표 모형

(Figure 5) A metric model of the type of social risk

5. 결 론

인공지능 로봇의 활용 범위가 작업 수행에서 일상생활 까지 확장되어 감에 따라 인공지능의 윤리 지표만으로는 인공지능 로봇의 위험성을 전반적으로 점검할 수가 없다.

인간과 상호작용이나 협업 중 인공지능 로봇으로 인해 인간이 다치거나 수집된 개인정보의 유출 및 차별 서비스 등 기술적 문제·윤리적 위험성 등 전체를 측정할 수 있는 지표가 있어야 한다. 이 연구에서는 인공지능 로봇의 기술 및 윤리적 문제 등 인간의 신체, 권리, 재산, 사회적 위험성을 측정할 검증된 지표를 개발하고자 하였다. 이를 위해 인공지능 로봇의 위험성 측정지표에 관한 델파이 연구에 참여했던 전문가 11명을 심층 면접하였다. 앞으로 산업, 공공

등 다양한 분야에서 인공지능 로봇이 도입될 수 있으므로 관련 직업 종사자 328명에게 현장성 설문조사를 하였다.

4단계 통계 분석을 통해 인공지능 로봇의 위험성 유형별 측정지표를 개발하였다. 이 연구에서 제안하는 측정지표가 인공지능 로봇을 안전하게 사용하는데 필요한 많은 연구의 초석이 되어 표준화된 지표, 정책 등에 활용되기를 기대한다.

참고문헌(Reference)

- [1] Park Eun-kyung, A robot that mistook humans for boxes, killed people, Hankookilbo, 2023. <https://v.daum.net/v/20231108180921195>
- [2] Korea Institute of Science and Technology Planning and Evaluation, Ministry of Science and Technology Information and Communication, “The future of social robots as a result of the 2019 Technology Impact Assessment.” pp.151, Dongjin Culture Co., Ltd., 2019.
- [3] IAAE, ETHICS GUIDELINES FOR TRUSTWORTHY AI, European Commission, 2020. <https://iaae.ai/research/?q=ytoyontzojeyoijrzx13b3jkc3r5cguio3m6mzoiywxsjtzojq6inbhz2uio2k6mtt9&bmode=view&idx=4384062&t=board>
- [4] Byeon Sun-yong, “A Study on the Ethical Guideline for artificial intelligence robots - Focusing on 4 principles of artificial intelligence robot ethics -,” Journal of Ethics Education Studies, Vol.47, pp.233-252, 2018. <https://www.dbpia.co.kr/Journal/articleDetail?nodeId=NODE10769080>
- [5] Ryu Ji-woong “A Study on Legal Problems of Artificial Intelligence (AI) Robot - Focusing on EU Legislative Trends of RoboLaw -,” Korean Society for Land Construction and Construction, Vol.78, pp.301-326. 2017. <https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE11471740>
- [6] Lee Hae-won, “Progress and Future Prospects of the EU AI Bill,” Journal of KISO, Vol.51, pp.34-37, 2023. <https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE11440658>

- [7] Japanese Cabinet Office, Human-centered AI social principles, 2019.
- [8] Ministry of Science and ICT, Korea Information and Communication Technology Association 2023 “A Guide to Reliable Artificial Intelligence Development - Self-Driving Field,” Jinhan M&B Co., Ltd., 2023.
- [9] C. H. Lee, S. Y. BYUN, B. J. Kim, H. J. Kim, H. C. Choi, Y. G. KIM and J. W. Kim. “A Study on the Development of Ethics Checklist for Developers of Healthcare AI Robots in Home,” Korean Ethical Society, vol.1, no.132, pp.263-280, 2021.
<https://doi.org/10.15801/je.1.132.202103.263>
- [10] S. D. Hong and Y. W. Yoo, “A Study on the Use and Risk of Artificial Intelligence (Focusing on the eproperty appraiser industry),” The Journal of the Korea Contents Association, Vol.22, no.7, pp.81-88. 2022.
<https://dx.doi.org/10.5392/JKCA.2022.22.07.081>
- [11] Torresen J., “A Review of Future and Ethical Perspectives of Robotics and AI,” Robot Learning and Evolution, Vol.4, 2018.
<https://doi.org/10.3389/frobot.2017.00075>
- [12] Jobin. A, Lenca, M., Vayena, E. “The global Landscape of AI ethics guidelines,” nature machine intelligence. pp.389-399, 2019.
<https://www.nature.com/articles/s42256-019-0088-2>
- [13] Hosseinzadeh, M. Sinopoli, B. and Bobick, A.F., “Toward Safe and Efficient Human - Robot Interaction via Behavior-Driven Danger Signaling,” IEEE Transactions on Control Systems Technology, vol.32, no.1, pp.214-224, 2024.
<https://doi.org/10.1109/TCST.2023.3305100>
- [14] W. Jeong., “Legal review on the use of artificial intelligence robots,” Kyung Hee University Law Research Institute, Vol.55, no.4, pp.189-221, 2020.
<https://www.dbpia.co.kr/Journal/articleDetail?nodeId=NODE10730413>
- [15] Guerra, A., Parisi, F. and Pi, D., “Liability for robots I: legal challenges,” Journal of Institutional Economics, Vol.18, no.3, pp.331-343, 2022.
<https://doi.org/10.1017/S1744137421000825>
- [16] S. H. Kim, “Artificial intelligence technology for robot-human, robot-robot, robot-equipment-human collaboration,” The Magazine of the IEEE, Vol.49, no.4, pp.17-24. 2022.
<https://www.dbpia.co.kr/Journal/articleDetail?nodeId=NODE11100751>
- [17] Huang, M. H, and Rust, R T., “Engaged to a Robot? The Role of AI in Service,” Journal of Service Research, Vol.24, no.1, pp.30-41, 2020.
<https://doi.org/10.1177/1094670520902266>
- [18] H. K. Song, and S. J. Ahn, “A study on the measurement indicators for risk prevention of AI robots,” Journal of Internet Computing and Services, Vol.24, no.3, pp.27-41. 2023.
<https://doi.org/10.7472/jksii.2023.24.3.27>

● 저 자 소 개 ●



송 현 경(Hyun-kyoung Song)

2010년 세종 사이버대학교 상담심리학과(학사)

2016년 한국외국어대학교 교육대학원 상담심리학과(교육학 석사)

2024년 성균관대학교 일반대학원 컴퓨터 교육학과 졸업(박사 학위)

2008년 7월~현재 경찰청 소속

관심 분야 : 인공지능 로봇, AI 윤리, 정보보안, 사이버범죄 예방, 학교폭력 예방, 데이터 분류 모델링

E-mail : onlyone_teacher@naver.com