

## 보험사 생성형 AI 도입에 따른 손해사정사의 역할변화

(The changing role of loss adjusters due to the introduction of generative AI by insurance companies)

조유미\*

Yumi, Cho

### <국문초록>

생성형 AI의 발전과 보험업계 도입 사례가 증가하고 있다. 이에 따라 본 연구는 생성형 AI 기술이 손해사정 업무의 어떤 부분을 대체하거나 보완하고 있는지, 그 역할이 어떻게 변화하고 있는지를 분석하고자 하였다. 그 결과 손해사정 업무에 있어서 AI 도입에 따라 첫째, 자동화된 보고서 생성이 가능해지며 데이터 분석 및 보고서 작성 자동화로 손해사정사는 데이터 검토 및 AI의 결과를 검증하는 역할로 이동할 가능성이 있다. 둘째, 이미지 기반 사고판별이 가능해지는 주로 교통사고 손해사정 업무에 활용된다. 셋째, 전문성 강화로 복잡한 클레임 처리, 약관의 모호성과 법적 분쟁 또는 해석차이 발생 시 상황별 판단, 피보험자와의 협상과 감정조절, 보험사와의 의견 조율, 사기 여부 판단 및 조사실행 등의 영역으로 확장 변화한다.

AI의 도입은 긍정적인 면과 함께 데이터와 알고리즘, 윤리적, 사회적 위험을 동반한다. AI 리스크는 AI가 개인과 기업에 미칠 잠재적 부정적 결과로 정의되며, 이를 보험사가 인수할 경우 손해사정 업무는 다음과 같은 역량을 요구받는

\* 장안대학교 경영과 교수 (yumidream@jangan.ac.kr)

이 논문은 장안대 2025년 연구비지원에 의해 연구되었음

다. AI 오류를 예견하거나 회피할 수 있는 가능성이 있었는가의 여부, 오류 발생 시 위험방지 조치 및 대처, AI 개발자와 사용자, 외부 공격자 사이의 과실의 출처와 증명, AI 오류로 인한 피해 발생 시 피해액 산정 등이 손해사정 업무에 포함된다. 또한 보험약관 상 AI 관련 손해가 계약상 담보대상인지 판단해야 하며 해당 사고가 우발적인 것인지 소프트웨어 오류가 면책사유인지 여부를 심사하고 이후 직간접 손해액을 산정해야 한다. 결론적으로 AI리스크에서 손해사정사는 단순 보상 평가자를 넘어서는 기술-법률-보험 통합 전문가의 역할을 요구 받는다. 이는 윤리적, 제도적 정당성을 입증하고 감정적, 기술적 판단과 함께 리스크 관리 및 정책개선 제안까지 포함한다.

※ 국문 주제어 : 생성형AI, 보험사AI도입, 손해사정사 역할변화, AI 리스크

## I. 서론

AI는 Turing(1950) 연구를 통해 최초로 정의되었으며, 연구자의 이름에서 유래한 튜링 테스트를 통해 인공지능 여부를 판단하였다. 튜링 테스트는 특정 질문에 대한 답변을 인간과 컴퓨터 모두에게 받고, 이후 답변이 인간의 답변인지 컴퓨터의 답변인지 구분 할 수 없다면 이를 인공지능이라고 판단했다. 초기 연구는 인간지능을 모방하여 인간을 대신하는 사고 및 패턴을 대신 할 수 있는 알고리즘을 만드는 것을 목표로 하였다. 이후 기술발전으로 신경망 이론, 자연어 처리, 음성처리, 이론 증명 등의 분야의 연구로 발전하여 딥러닝(Deep Learning) 알고리즘<sup>1)</sup>을 거쳤다. 최근에는 생성형 AI의 등장으로 사고력과 학습력을 획기적으로 향상시켰다.

생성형 AI는 인간 피드백에 따른 강화학습(Reinforcement Learning Through Human Feedback: RLHF)<sup>2)</sup>을 통해 방법을 통해 개발자가 원하는 대로 데이터가 생성되도록 설계되었다. 창의적인 이미지, 영상, 음악, 텍스트를 생성하고 기존 데이터 심화학습을 통해 시, 소설, 기사를 작성할 수 있다. AI 개발 초기에는 인간을 모방하는 것에서 그쳤으나 이후 스스로 생각하고 학습하게 되었으며 최근에는 자기계발이 가능해지고 있다. 이에 따라 산업과 사회 전반에서 AI의 활용은 광범위해졌으며 그 중 금융보험 산업이 가장 높은 AI 채택지수(Adoption Index)를 보이고 있다(Human Centered AI, 2023)<sup>3)</sup>.

보험산업에서 생성형 AI의 규모는 2025년 현재 837.8(million\$)에서 2030년 3040(million\$)로 확대, 2032년에는 5543.1(million\$)로 확대될 것으로 예측되고 있

- 1) 비선형 변환기법의 조합을 통해 수많은 데이터 속에서 핵심적인 내용 또는 특징을 단시간에 요약해주는 기계학습(Machine Learning) 알고리즘 집합이다.
- 2) RLHF(인간 피드백을 통한 강화 학습)는 머신러닝 모델을 인간의 선호도와 피드백에 맞춰 최적화하는 기법이다. 이 과정은 먼저 인간의 직접적인 평가를 바탕으로 보상 모델을 학습하고, 이 보상 모델을 이용해 강화 학습을 통해 AI 에이전트의 성능을 향상시키는 방식으로 이루어진다.(Lee et. al, 2024)
- 3) 제품 및 서비스 개발 분야 AI 채택지수는 모든 산업 평균 10%로 법률 및 경영 8%, 소매업 4%, 의료업 4%, 통신업 7%, 금융보험업31%이며 시스템 운영 분야 AI 채택지수는 모든 산업 평균 19%로 법률 및 경영 20%, 소매업 31%, 의료업 12%, 통신업 21%, 금융보험업24%.

다. 보험사는 고객 니즈에 맞는 상품개발, 손해율, 언더라이팅, 상품판매, 보험금지급, 고객관리 등의 보험업 모든 분야에서 금융데이터 뿐만 아니라 비금융데이터를 방대하게 활용(이효섭, 2023)하고 있다. 또한 경제 주체들 간의 정보 비대칭으로 인한 갈등과 신뢰부족이 여전히 완전 해소되지 않은 상태를 지속하며 민원과 법적분쟁이 끊이지 않고 있다. 이러한 문제들을 타계할 대안으로 생성형 AI 활용 확대는 보험사의 수익성과 성장성을 제고한다.

보험사의 생성형 AI기술 도입에 따라 가장 중요한 역할변화가 예상되는 분야 가운데 하나는 손해사정이다. AI를 활용하는 기업이나 산업에서 손해 발생 시 해당 피해가 AI로 인한 것인지 인간에 의한 것인지 여부를 판단해야하며, 손해사정사가 피해의 규모와 범위를 어떻게 한정하는가에 따라 상품의 성질이 확연히 달라질 가능성이 있기 때문이다. 또한 생성형 AI를 손해사정 업무에 활용할 경우 그 결과의 차이와 발전이 예상된다. 따라서 향후 보험사가 인수하게 될 AI 관련 리스크에 대한 위험평가 방법과 범위를 포함한 손해사정에 대한 연구가 필요하다.

본 연구는 생성형 AI 도입에 따른 손해사정사의 역할 변화<sup>4)</sup>, AI 기반 손해사정의 법적 책임에 관한 연구<sup>5)</sup>이다. 생성형 AI의 도입은 손해사정사의 역할을 평가자에서 AI 활용 전문가, 데이터 검증자, 고객 신뢰 구축자로 변화시키고 있다. 이에 따라 AI의 정확성과 공정성 연구, 손해사정사와 AI의 협업 모델 개발, 윤리적·법적 문제 해결은 중요한 연구 과제이다. 앞으로 보험업계와 학계는 AI가 손해사정의 효율성을 높이면서도 공정성을 유지할 수 있도록 지속적인 연구와 제도적 정비가 필요하다.

4) 데이터 분석 및 보고서 작성 자동화로 손해사정사는 데이터 검토 및 AI의 결과를 검증하는 역할로 이동할 가능성이 있으며 AI가 산출한 결과를 해석하고 예외적인 사례를 판단하는 데 집중하게 된다. 간단한 손해사정 업무는 AI가 수행하고, 손해사정사는 고객과 직접 상담하며 복잡한 사례(예: 사기 가능성이 있는 사건, 중대한 법적 분쟁)를 처리하는 역할을 맡는다. AI가 생성한 손해 평가 결과의 정확성을 검토하고, 데이터 윤리 및 개인정보 보호 관리, 고객 신뢰 구축 및 대인관계 능력 강화, AI의 한계를 보완하는 역할을 수행해야 한다.

5) AI가 잘못된 손해 평가를 했을 때, 책임이 보험사, AI 개발사, 손해사정사 중 누구에게 있는지에 대한 법적 연구가 필요하다. 특히, AI가 클라우드 기반으로 데이터를 처리하는 경우 보안 취약점이 없는지 점검하는 연구가 필요하다.

## II. 전통적 손해사정사의 역할과 AI 도입 후 역할변화

### 1. 전통적 역할

보험업법 제185조에 의거 손해사정은 보험회사가 손해사정사를 고용하여 보험 사고에 따른 손해액 및 보험금의 사정업무를 처리하는 것으로 정의하고 있다. 또한 보험업법 제188조에 의거 손해사정사는 손해발생사실의 확인, 보험약관 및 관계법규 적용의 적정성 판단, 손해액 및 보험금 사정, 해당 업무에 대한 서류작성 및 제출대행, 업무와 관련된 보험회사에 대한 의견의 진술 등의 업무를 수행한다. 따라서 손해사정사는 보험사와 보험가입자(피보험자) 사이에서 보험금 지급 여부 및 지급액을 공정하게 판단하는 역할을 한다. 보험금이 청구 되었을 때 사고의 원인, 손해의 정도, 보험 계약 내용 등을 종합적으로 검토하여 보험금 산정에 필요한 판단을 한다. 이는 금융위원회에서 인정하는 국가자격증을 통해 활동이 이루어지며 보험회사 소속의 전속 손해사정사, 독립적인 개인 손해사정사 또는 손해사정 법인 소속으로 활동한다. 보험손해사정사의 핵심역량은 협상력, 회복탄력성, 고객지향성 및 업무공정성, 문서 작성 능력이며(임채훈, 김주일, 2023), 활동 내용은 아래와 같다.

첫째, 실제 사고가 발생하였는지 확인하고 그 원인과 경과를 명확히 파악하는 것을 목적으로 사고발생 경위 조사, 피해 현장 조사 및 자료 수집<sup>6)</sup>, 관련자 인터뷰 진행 등을 통해 사고조사를 한다.

둘째, 실제 경제적 손실의 규모를 객관적으로 판단한다. 재산 및 신체에 대한 손해규모 계산을 위해 차량 수리 견적서, 치료비, 위자료, 일실수입(수입손실), 휴업손해 등의 자료와 필요시 의사, 정비사, 건축 전문가 등의 협업을 통해 손실 규모에 대한 적절한 판단을 내린다.

셋째, 해당 사고의 보험금 지급 대상 여부를 판단하는 것을 목적으로 보장범위, 면책사유, 가입금액 및 특약사항 등을 확인하여 보험 계약과 약관 내용에 따라 보

6) 사진, 진술, 경찰 보고서, 병원 진단서, 영수증 등의 보험금 청구에 필요한 서류

상 가능 여부 판단 및 보험사기 여부를 판단한다. 이때 보험업법, 민법, 상법 등 관련법령에 의해 판단하며 보험사기 가능성이 있을 경우 형사고발도 가능하다.

넷째, 객관적이고 적절한 보험금을 산정하여 보험사에 보고하고 추천의견을 제시한다. 실제 손해액과 보상 한도를 비교하여 보험금을 산출하며 자기부담금 차감 등의 보험약관에 따른 조정을 하여 사고개요, 조사결과, 산정 근거, 지급 권고액 등을 명시해 보험사에 최종 보고서를 제출한다.

마지막으로 보험금 지급과 관련된 보험소비자와 보험사 간 분쟁 발생 시 전문가로서의 의견을 제시하고 법적 소송 시 증언 또는 감정서를 제공한다.

## 2. 손해사정 업무의 변화 및 미래역할

대형 언어모델, 멀티모달 모델 등의 생성형 AI의 도입은 손해사정사의 업무 중 많은 부분을 자동화 또는 보조할 수 있으며 그 결과로 손해사정사의 역할은 전통적 수행자에서 감독자, 고도 조정자, 전략적 분석가의 방향으로 이동할 가능성이 높다. 이는 반복적 문서 중심 업무의 자동화, 리스크 평가, 보험금 누수(leakage) 감소, 관계 교섭, 조직 내 업무 재설계, 윤리 설계 등의 역할 포함한다.

손해사정 업무에 있어서 AI 도입에 따른 실질적 변화는 크게 3가지 정도로 구분 가능하며 결론적으로 AI는 손해사정사가 수행하는 전문적인 업무를 돕거나 협업하는 형태가 될 것으로 예상된다.

첫째, 자동화된 보고서 생성이 가능하다. 데이터 분석 및 보고서 작성 자동화로 손해사정사는 데이터 검토 및 AI의 결과를 검증하는 역할로 이동할 가능성이 있으며 AI가 산출한 결과를 해석하고 예외적인 사례를 판단하는 데 집중하게 된다. 간단한 손해사정 업무는 AI가 수행하고, 손해사정사는 고객과 직접 상담하며 복잡한 사례, 예를 들어 사기 가능성이 있는 사건, 중대한 법적 분쟁을 처리하는 역할을 맡는다. AI가 생성한 손해 평가 결과의 정확성을 검토하고, 데이터 윤리 및 개인정보 보호 관리, 고객 신뢰 구축 및 대인관계 능력 강화 등 AI의 한계를 보완하는 역할을 수행하게 된다(조유미, 2025).

AI는 고객과의 통화기록 전사, 후속메일 초안 작성, 내부 프로세스 설계 등의 역할을 한다. 다만 문서생성과정에서 오류, 맥락이해부족으로 인한 환각위험(hallucination risk)<sup>7)</sup>이 존재하므로 손해사정사가 이를 관리감독 해야 한다. 또한 손해사정 과정에서 보험계약 조건 미준수, 과잉지급, 잘못된 보상 등이 있을 경우 AI는 과거 사례 패턴을 학습하여 비정상 영역(leakage)<sup>8)</sup>을 감지하고 조정 제안을 할 수 있다(Bain & company, 2024).

둘째, 이미지 기반 사고판별(Image-Based Accident Assessment)이 가능하다. 이는 주로 교통사고 손해사정 업무에 활용된다. 이는 차량 파손 사진, 도로 상황 사진 등의 사고현장 이미지 데이터를 인공지능 알고리즘, 특히 컴퓨터 비전 기술로 분석하여 손해 사정에 필요한 정보를 도출하는 과정을 의미한다. 분석대상은 차량의 파손부위, 파손정도, 위치, 손상유형(굽힘, 찌그러짐, 파손 등)이며 이에 사용되는 기술은 CNN(Convolutional Neural Networks)<sup>9)</sup>, Object Detection<sup>10)</sup> 등이 대표적이며 이미 현대 오토에버, 테슬라 등이 AI 플랫폼 기술을 도입하여 사용하고 있다.

CNN 기반 이미지 분류 모델은 차량의 앞 범퍼, 측면, 뒷부분 등 손상위치를 자동으로 분류할 수 있으며, Object Detection은 충격지점을 박스(Bounding Box)로 표시하고 손상부위를 시각적으로 강조하여 보여준다. 이러한 시각자료를 토대로 AI는 손상의 면적, 깊이 등을 분석하여 수리비용 예측모델과 연계할 수 있으며 일부 시스템은 과거 유사 사례 데이터를 학습하여 수리에 필요한 부품 또는 클레임 부품식별(task previously manual)과 인건비까지 자동으로 산출한다.

7) 사실이 아닌 정보생성

8) 정책 가이드라인에서 벗어나거나 공급망 문제로 인해 예상치 못한 비용이 발생할 때 지불된 금액과 지불해야할 금액의 차이

9) 합성곱 신경망(CNN)은 특히 이미지와 같은 격자형 데이터를 처리하는 데 탁월한 딥 러닝 신경망 유형으로, 특수 합성 계층과 필터를 사용하여 패턴과 기능을 자동으로 감지한다. 이러한 네트워크는 합성곱 계층, 풀링 계층, 완전 연결 계층을 포함한 여러 계층으로 구성되어 있으며, 기본적인 에지에서 고수준 객체 부분까지 점점 더 복잡한 기능을 학습하며 이미지 인식, 자연어 처리, 자율 주행 자동차 기술과 같은 작업에 널리 사용된다.

10) 객체 감지(Object Detection)는 컴퓨터 비전 분야에서 이미지나 비디오 내의 객체들을 찾아내고, 각 객체의 위치와 종류를 분류하는 기술이다. 이는 딥러닝이나 머신러닝 알고리즘을 활용하여, 이미지에서 인간처럼 객체들을 인식하고 그 주변에 경계 상자를 그려서 위치와 종류를 파악하는 컴퓨터의 지능을 재현하는 것을 목표로 한다.

이러한 AI 기술을 손해사정 업무에 도입하게 되면 기존에 손해사정사가 현장을 방문, 수기로 문서작성, 사진분석, 견적서 비교 등의 업무를 하던 것에서 이후 이미지 업로드를 통해 AI 자동분석을 통해 손해추정 보고서를 자동으로 생성 받을 수 있으며 이에 따라 업무 속도가 향상되고 표준화된 평가를 제공할 수 있다.

셋째, 손해사정사의 전문성이 강화된다. 우리나라는 경제규모 확대와 담보 범위 다양화, 시장 대형화 등의 이유로 손해발생 원인, 면책과 부책의 판단, 손해액과 보험금액 결정 등이 복잡해졌다. 이에 따라 손해사정 업무에 있어 기술 및 전문성의 고도화가 요구되고 해당 업무를 이행하는 손해사정사에게 이러한 요구가 뒤따른다. 이러한 이유로 손해사정사의 전문성 제고를 위해 시험과목 및 자격인증 제도를 강화하고 면허갱신제도, 보수교육 의무화를 추진해야 한다는 학계의 주장이 대두되고 있다(이상실 2023, 백주민 2024, 김명규·마승렬 2025). 이와 더불어 본 연구는 AI 역량을 강화하기 위한 손해사정사 재교육이 필요하다고 사료한다.

손해사정사 전문성이 필요한 핵심영역은 다수의 이해관계자 관여 또는 책임소재 불분명한 손해사건에서의 심층적 판단을 통한 복잡한 클레임 처리, 약관의 모호성과 법적 분쟁 또는 해석차이 발생 시 상황별 판단, 피보험자와의 협상과 감정조절, 보험사와의 의견 조율, 사기 여부 판단 및 조사실행 등의 영역이다. 이는 윤리적, 제도적 정당성을 입증하고 감정적, 기술적 판단과 함께 리스크관리 및 정책개선 제안까지 포함한다.

### III. AI 리스크

#### 1. AI 리스크 생성 및 확대

생성형 AI의 활용은 이에 따르는 리스크를 동반한다. AI 리스크는 AI가 개인과 기업에 미칠 잠재적 부정적 결과로 정의되며, 수행, 보안, 통제, 사회, 경제, 윤리적 리스크로 구분된다. 수행 리스크는 예측오류, 편향, 알고리즘 불투명성, 알고



리즘 설명 불가능, 결과의 불안정성이며, 보안 리스크는 사이버 보안침해, 개인 정보 침해, 오픈소스 소프트웨어 취약성, 통제 리스크는 AI 악용 가능성, 악의적 사용 AI 통제능력 부족, 사회 리스크는 자율 무기 양산, 계층간 불평등 초래, 경제 리스크는 AI 인간 업무 대체, 승자독식, 배상책임, 평판리스크 등이다. 마지막으로 윤리 리스크는 윤리적 가치 결여, 가치 및 목적 협의 부족 등이다 (Brookings Institution, 2019). 본 연구는 세분화되어 있는 AI리스크를 기술, 윤리, 사회, 법적 측면에서 구분하고자 한다.

첫 번째 기술적 측면에서 AI 리스크는 데이터와 알고리즘에 기인한 예측오류 가능성이 있다. 구성요소에 따라서 AI 리스크는 데이터 리스크와 알고리즘 리스크로 나눌 수 있다. 데이터 리스크는 과거 데이터 활용에 따른 자료편향(selection bias), 데이터 신뢰 가능성에 따른 불투명성, 과거 데이터의 전체 손실, 과거 데이터 부분손실로 인한 편향성과 과소 또는 과대 추정 등의 문제이다. 한편 알고리즘 리스크는 복잡성, 복잡성으로 인한 자의적 해석, 패턴 해석 불가, 추정오류 및 거짓추정 생성, 추정오류 확인불가 등의 위험이다.

데이터의 외연적 확장과 학습 알고리즘의 발달에 의해 AI가 급속히 성장한다는 것은 새로운 데이터를 활용하면서 발생하는 리스크와 복잡한 알고리즘 사용으로 인한 오류 및 해석의 문제 발생을 야기한다(Kautz, 2022). 사용 가능한 데이터의 급격한 증가는 찾고자 하는 모듈과 사용된 데이터의 정보적 연관성을 보장할 수 없는 이유로 유용성 저하가 우려된다. 더욱이 비정형 데이터의 경우 일관된 형태로 가공하는 과정에서 정보의 변형, 손실 발생 가능성이 있다(Loughran & McDonald, 2011). 또한 AI 모델의 입력 데이터가 약간만 달라져도 전혀 다른 결과를 도출하며(MIT CSAIL, 2023) 이를 적대적 공격(adversarial attack)으로 명명하고 있다. 이는 AI 시스템이 의도와 다르게 작동하거나 잘못된 판단을 내리는 것으로 예를 들어 자율주행 자동차의 비정상 작동으로 인한 사고발생 가능성 등을 의미한다.

또한 알고리즘 리스크는 주로 학습과정의 불투명성과 복잡성에 기인한다. 이러한 불투명성을 해결하기 위해 XAI<sup>1)</sup>의 국내도입을 주장하는 연구(조유미,

2025)가 있으나 현업 전체에 활용되기에는 시기적, 법률적 문제가 남아있다. 아울러 머신러닝 기반 AI는 대량의 모수(parameter) 데이터 학습으로 규칙을 형성하고 효용성 판별로 정량적 목적함수(objective function)를 사용하는 일차원적 특징을 가짐으로서 현실 단순화 오류의 가능성을 지닌다. 이를 해결하기 위해 알고리즘에 기반한 의사결정이 사회에 미치는 외부성 고려 주장(Korinek & Balwit, 2022)이 있으나 아직까지는 현실적 시스템 문제가 있다.

AI 리스크는 AI 기반 의사결정에 따른 문제가 발생할 경우 그 책임소재 불확실성을 가진다. 데이터 측면에서 예측모형은 과거 데이터에 기반하여 학습을 진행하게 되고 이는 예측 결과가 과거 데이터 편향적 분포의 성향을 가질 수 있음을 의미한다. 따라서 가용한 데이터의 범주 내에 존재하지 않는 사건<sup>12)</sup>은 기존 예측치의 정합성 보장이 어렵다. 또한 복잡한 알고리즘을 통해 생성된 결과는 개발자 및 이용자의 직관적 해석을 어렵게 한다. 이는 알고리즘의 설계, 오류와 시정의 범위, 알고리즘이 개인에게 미치는 영향의 정도 등의 투명성 논란을 야기하며 결국 문제발생시의 원인 해명과 법적, 도의적 책임판단의 문제를 발생시킨다(조유미, 2025).

두 번째는 윤리적 리스크이다. 대규모 AI모델이 민감한 개인정보를 학습하거나 노출시켜 생기는 위험성이다. Meta AI<sup>13)</sup>의 LLama 모델<sup>14)</sup>이 Reddit/Stackoverflow의 사용자 데이터를 무단으로 사용했다는 논란이 있는가 하면 Chat GTP가 사용자의 대화내용을 무단으로 학습한 일로 인해 이탈리아 정부는 2023년 한시적으로 GTP사용을 금지하기도 하였다. EU AI ACT(2024)<sup>15)</sup>는 고위험 AI에 대해 법적 책

11) XAI(Explainable AI)는 블랙박스의 구조와 개별적인 결정의 인과관계를 보다 투명하고 신뢰할 수 있도록 하려는 활동으로 AI를 활용하여 데이터를 분석 할 경우 인사이트 추출 및 해석에 있어 설명 가능한 상태를 만든다.

12) 예를 들어 COVID-19, 우크라이나 러시아 전쟁, 대통령 탄핵 등

13) Meta AI는 Meta사가 개발한 가상 비서의 이름이기도 하며, 현재는 Meta의 소셜 네트워킹 제품에 챗봇으로 통합되어 있다.

14) Meta AI 가 2023년 2월부터 출시한 대규모 언어 모델 (LLM) 제품군으로 최신 버전은 2025년 4월에 출시된 Llama 4이다.

15) EU AI법은 인공지능에 대한 최초의 글로벌 규제 체계로, 유럽 연합 내 AI 시스템의 개발, 사용 및 배포에 대한 위험 기반 규칙을 수립하여, 2024년 8월 1일부터 시행되었다. 이 법은 AI의 신뢰성을 보장하고, 기본권을 보호하며, 혁신을 촉진하고, EU 회원국 전반에 걸쳐 통일되고 조

임소재를 명확히 규정할 것을 요구하고 있으나 실제로 소재의 분명한 주체를 구분하는 것은 쉽지 않다. 아울러 책임이 AI에게 있다고 하더라도 인격을 갖추고 있지 않는 존재에게 윤리적 책임을 묻는 것에는 한계가 있다.

개인정보 유출 및 인간중심 판단 미흡 등의 문제가 발생할 가능성이 다분하다. AI 기반의 챗봇 사용으로 금융소비자의 대화 메시지가 노출되거나 챗봇 서비스에서 금융계좌 해킹, 알고리즘 오류, 개발자의 악의적 활용, 특정 금융상품의 마케팅 악용 될 가능성이 있다. 또한 자본력이 갖춘 대형 금융사가 우수한 생성형 AI를 개발 또는 도입해 지배력을 강화하여 중소형 금융회사들의 수익성 급락의 우려도 있다.

또한 데이터 가공의 측면에서는 구분 및 분리로 보아 법적인 문제로 대두되지 않는다고 하더라도 윤리적 측면에서 비공식적인 인종차별, 지역차별, 성차별, 기호차별 등의 각종 차별이 문제가 된다. ProPublica(2016)의 보고서는 범죄예측 AI 시스템이 흑인에게 불리한 판단을 내리고 있다고 지적하기도 하였다.

세 번째는 사회적 리스크이다. Goldman Sachs<sup>16)</sup>의 2023년 Economics Research는 자동화된 시스템으로 인해 특정 직종이 대략으로 사라질 가능성에대한 우려를 언급하고 있다. 생성형 AI는 노동 비용 절감, 생산성 향상, 그리고 인간과 기계 간의 소통 장벽을 허물어 거시경제에 큰 영향을 미칠 수 있는 주요 발전요소로 작용할 수 있다. 그러나 이는 전 세계적으로 약 3억 개의 정규직 일자리를 자동화에 노출시킬 수 있으며, 노동 시장에 상당한 혼란을 야기할 수 있다(Briggs & Kodnani, 2023). 특히 회계, 고객센터, 법률보조, 번역 등의 직업군이 고위험군이며, 나아가 인간의 통제를 벗어난 초지능 AI(AI Super-intelligence)는 장기적으로 인류에 위협이 될 위험이 있다. 또한 딥페이크(Deep fake) 등을 이용하여 AI가 허위정보를 생성하거나 진짜처럼 보이게 만들 수 있다. 2024년 미국 대선을 앞두고 AI가 만든 가짜 바이든 음성이 확산되는 사건 등과 같이 허위사실 유포 및

화로운 법적 환경을 조성하는 것을 목표로 하며 용납할 수 없는 위험을 수반하는 특정 AI 관행을 금지하고, 고위험 시스템에 엄격한 요건을 부과하며, AI의 투명성과 안전성을 장려한다.

16) 글로벌 투자 은행, 증권, 자산 및 재산 관리 회사

증거조작의 가능성이 있다. 마지막으로 법적 및 규제 리스크가 있다. 적절한 법적 프레임이 부재하거나 과도한 규제로 혁신이 저해될 소지가 있으며, 앞서 언급한 바와 같이 불분명한 책임소재는 법적 판단을 저해한다.

## 2. AI 리스크 인수

AI 도입으로 발생한 여러 가지 리스크를 보험사가 인수하게 되면 보험사 입장에서 새로운 새롭고 거대한 시장이 될 가능성이 크다. AI관련 비즈니스 리스크에 대한 보험료가 연평균 80%의 증가율을 보이며 2032년까지 한화 약 6조 1,608억원에 이를 것으로 전망되고 있다(Deloitte, 2024). 현재 AI관련된 보장을 하는 보험은 사이버 리스크와 같이 대체로 보안 리스크에 중점을 두고 개발(김윤진, 2021)되고 있기 때문에 향후 AI 활용확대로 인해 AI 사용에 따르는 신체상해, 물적 손해, 평판 리스크, 수행 리스크 등의 증가에 따른 대응이 요구된다. AI 활용으로 발생 가능한 다양한 위험을 보장하기 위해서는 AI 모델에 대한 정확한 정보와 성능에 대한 데이터를 확보하는 것이 중요하며, 이에 따라 보험사의 위험평가 역량도 키워야 할 것이며 보험사가 새로운 리스크를 인수한다면 이에 따르는 손해사정도 반드시 뒤따를 것이다. 이를 기회로 만들 수 있는 손해사정업계의 노력이 절실하다.

AI 리스크는 보험사 입장에서 몇 가지 측면으로 구분하여 인수 가능하다. 우선 AI시스템 오류 관련 배상책임보험으로 AI의 오작동으로 인해 잘못된 의사결정을 하여 고객사에게 손해가 발생했을 경우이다. 예를 들어 병원용 AI진단 시스템이 오진을 하여 환자에게 피해가 발생한 경우에 손해배상 청구 등을 기술 전문배상 책임보험(Tech E&O)의 형태로 인수 가능하다. 또한 AI 시스템이 사이버 공격을 받거나 프라이버시 침해가 발생했을 경우 사이버 보안 진단 기반 인수 심사가 가능하며, 자율주행자동차 및 로봇 등과 같이 AI가 탑재된 제품에서 사고가 발생할 경우 제조사 책임을 담보할 수 있다. 아울러 금융회사의 자동거래 AI 중단으로 인해 수익손실이 발생하거나 e-커머스 플랫폼의 추천 AI 시스템 장애로 인한 매

출이 급감하는 등의 경우 전통적 업무중단보험(Business Interruption)과 연계하거나 복구계획(BCP), 백업체계 등을 인수 심사에 반영하는 등 AI시스템의 장애 또는 다운타임(down time)으로 인해 기업에게 발생한 손실을 인수할 수 있다.

보험사가 AI 리스크를 인수하기 위해서는 결정과정 설명가능성에 근거한 알고리즘의 투명성, 학습 데이터의 차별 요소 포함여부에 따른 데이터 품질과 편향성(bias), 사업모델의 AI 의존도 등을 우선 평가해야 한다. 이때 AI 사고에 대한 보험 손해를 데이터가 부족하여 리스크를 정량화하기 어렵거나 법적 규제환경과 보험 약관 설계 변동성이 존재하는 경우 이를 반영해야 한다. 결과적으로 보험사는 AI 리스크를 기술적, 윤리적, 운영적 리스크로 구분하여 관리해야 하며 AI 시스템 보유 기업에 대해 기술, 법무, 보안 실사를 진행해야 한다. 또한 데이터 전문가를 포함한 AI/Tech 전담 언더라이팅팀을 운영하고 대형 리스크 분산을 위해 재보험사와의 협업이 요구되기도 한다.

보험사의 AI관련 리스크 인수는 손해사정업 측면에서 기회가 될 수 있다. 우선 사고원인 분석과 책임소재 판단에 있어서 손해사정사의 역할이 필요하다. AI 리스크는 인간의 사고 또는 손해 판단을 보조하는 AI의 역할 수행에 있어서 AI의 오류가 발생할 경우 이에 따른 손해배상에 대한 책임소재에 대한 내용이 주를 이룬다. 이는 손해사정 업무에 활용되는 AI뿐만 아니라 사회 전반에 걸쳐 활용되는 AI가 잘못된 학습으로 인한 오류 또는 학습 데이터의 저작권 및 소유권 침해 등의 문제가 발생했을 경우 이에 대한 손해이다. AI 오류를 예견하거나 회피할 수 있는 가능성이 있었는가의 여부, 오류 발생시 위험방지 조치 및 대처, AI 개발자와 사용자, 외부 공격자 사이의 과실의 출처와 증명, AI 오류로 인한 피해 발생 시 피해액 산정 등 손해사정 업무에 해당한다. 사회 전반에서 활용되는 AI가 오류를 발생시키거나 반드시 오류가 아니더라도 법적 분쟁의 소지가 될 가능성이 다분하며 공동 책임 여부 판단이 필요하기도 하다.

또한 보험약관 상 AI 관련 손해가 계약상 담보대상인지 판단해야 하며 해당 사고가 우발적인 것인지 소프트웨어 오류가 면책사유인지 여부를 심사하고 이후 직간접 손해액을 산정해야 한다.

AI 사고의 경우 손해사정상 특수성이 존재한다. 전통적으로는 가시적이고 물리적인 사고판단이었던 AI리스크는 보이지 않는 알고리즘이나 데이터 오류에 의한 사고이며, 특정 개인이나 법인과 같이 명확한 사고 책임자가 있었다면 AI리스크에서는 복수의 이해관계자와 더 많은 사고 책임자가 존재한다. 기존에 보험약관상 명확한 분류기준을 가지고 사정하던 방식에서 새로운 유형 약관으로 인한 복잡성이 초래하고 법률, 회계 중심 사정과 더불어 기술, IT보안, 알고리즘에 대한 이해를 필요로 한다. 결론적으로 AI리스크에서 손해사정사는 단순 보상평가자를 넘어서는 기술-법률-보험 통합 전문가의 역할을 요구 받는다.

#### IV. 결론 및 향후연구과제

본 연구는 보험산업의 디지털 전환속에서 손해사정사의 역할을 정립하고 AI와 인간 전문가의 효율적 협업모델을 제시하고자 하였다. 또한 정책 및 교육 프로그램 마련을 위한 기초자료를 제공하고자 한다.

손해사정 업무에 있어서 AI 도입에 따른 실질적 변화는 아래와 같다.

첫째, 자동화된 보고서 생성이 가능하다. 데이터 분석 및 보고서 작성 자동화로 손해사정사는 데이터 검토 및 AI의 결과를 검증하는 역할로 이동할 가능성이 있다.

둘째, 이미지 기반 사고판별(Image-Based Accident Assessment)이 가능하다. 이는 주로 교통사고 손해사정 업무에 활용된다.

셋째, 전문성 강화로 복잡한 클레임 처리, 약관의 모호성과 법적 분쟁 또는 해석차이 발생 시 상황별 판단, 피보험자와의 협상과 감정조절, 보험사와의 의견조율, 사기 여부 판단 및 조사실행 등의 영역으로 확장 변화한다.

AI의 도입은 긍정적인 면과 함께 위험을 동반한다. AI 리스크는 AI가 개인과 기업에 미칠 잠재적 부정적 결과로 정의되며, 본 연구는 데이터 및 알고리즘 리스크, 윤리적 리스크, 사회적 리스크로 구분하였다.

첫째, AI는 데이터와 알고리즘에 기인한 예측오류 가능성이 있다. 알고리즘 리스크는 주로 학습과정의 불투명성과 복잡성에 기인하며 AI 기반 의사결정에 따른 문제가 발생할 경우 그 책임소재 불확실성을 가진다.

둘째, 윤리적 리스크는 대규모 AI모델이 민감한 개인정보를 학습하거나 노출시켜 생기는 위험성이다. 비공식적인 인종차별, 지역차별, 성차별, 기호차별 등의 각종 차별이 문제가 여기에 포함된다.

세 번째는 사회적 리스크이다. AI는 전 세계적으로 약 3억 개의 정규직 일자리를 자동화에 노출시킬 수 있으며, 노동 시장에 상당한 혼란을 야기할 수 있으며, 딥페이크(Deep fake) 등을 이용하여 AI가 허위정보를 생성하거나 진짜처럼 보이게 만들 수 있다.

위와 같은 AI리스크를 보험사가 인수할 경우 손해사정 업무는 다음과 같은 역할을 요구받는다. AI 오류를 예견하거나 회피할 수 있는 가능성이 있었는가의 여부, 오류 발생시 위험방지 조치 및 대처, AI 개발자와 사용자, 외부 공격자 사이의 과실의 출처와 증명, AI 오류로 인한 피해 발생 시 피해액 산정 등이 손해사정 업무에 포함된다. 또한 보험약관 상 AI 관련 손해가 계약상 담보대상인지 판단해야 하며 해당 사고가 우발적인 것인지 소프트웨어 오류가 면책사유인지 여부를 심사하고 이후 직간접 손해액을 산정해야 한다. 결론적으로 AI리스크에서 손해사정사는 단순 보상 평가자를 넘어서는 기술-법률-보험 통합 전문가의 역할을 요구 받는다. 이는 윤리적, 제도적 정당성을 입증하고 감정적, 기술적 판단과 함께 리스크관리 및 정책개선 제안까지 포함한다.

2025년 9월 현재 보험연수원은 집합교육 과정으로 AI전략, 데이터 기획 및 분석, 인슈어테크, 스마트 워크 등의 교육과정을 초급, 중급, 고급으로 분류 운영 중이다. 이러한 연수 및 교육과 더불어 보험사에게 인수된 AI 리스크의 손해사정을 담당하는 전문인력으로서 손해사정사 실무교육 커리큘럼 개발 등에 대한 향후연구가 필요하다.

예를 들어 기초과정으로 AI, 머신러닝과 딥러닝의 차이, 자율주행 및 의료 AI 등 주요분야 소개를 내용으로 하는 AI기술개요, AI 시스템의 구조와 작동원리,

Tech E&O, 사이버 보험, 상품배상보험의 담보범위와 면책조항 등의 AI관련 보험상품에 대한 이해, 손해사정 절차, 관련 법률 및 규제, 디지털 포렌식 등을 교육한다. 나아가 심화과정으로는 매출손실, 브랜드 훼손 등의 무형손해 산정, 자율주행 AI 사고 종합분석, AI와 연계된 약관조항 분석, 약관 분쟁 사례 실무적용 방법, 실제사례 시뮬레이션 평가 보고서 작성 등을 골자로 한다.

변화된 환경에서 변화된 역량을 요구받는 손해사정업계 및 학계의 구체적이고 실무적인 AI 재교육과 이를 활용한 업무단계 체계화의 노력, 이를 위한 관련 금융감독기관 및 정부의 전폭적인 지원을 기대한다.



## 〈참고문헌〉

- 고상원, “신뢰할 수 있는 인공지능을 위한 규제 동향,” 경제인문사회연구회, 2023 Global Issue Brief, 7, 2023
- 김명규, 마승렬, “손해사정제도의 개선방안과 정책적 과제,” 한국손해사정학회, 손해사정연구 17(1), 2025, 29-56.
- 김윤진, “AI 리스크와 전용보험의 필요성,” KIRI 리포트, 보험연구원, 2021
- 딜로이트, “블루오션으로 떠오른 AI 리스크 보험 산업,” 2024.10.07
- 마승렬, 김명규, “위탁손해사정 보수의 개선방향,” 한국손해사정학회, 손해사정연구, 24(0), 2023, 27-50.
- 백주민, “손해사정 제도의 실태분석 및 개선방안에 대한 연구,” 한국테러학회보, 17(1), 2024, 53-64.
- 이상실, “손해사정사 자격제도와 직업의 자유에 관한 연구,” 박사학위논문, 동국대학교 일반대학원. 서울, 2023
- 이효섭, “생성형 AI가 금융산업에 미치는 영향-혁신과 리스크 요인,” 글로벌금융리뷰, 4(1), 2023, 91-121.
- 임채훈, 김주일, “보험손해사정사의 협상역량척도 개발에 관한 연구,” 분쟁해결연구센터, 분쟁해결연구 21(2), 2023, 63-91.
- 조유미, “보험사 생성형 AI 언더라이팅 활용 문제점과 해결방안 모색,” 한국손해사정학회, 손해사정연구 17(1), 2025, 5-28.
- Brent Winslow, Ozlem Ozmen Garibay, “Human-Centered AI,” Human-Computer Interaction in Intelligent Environments, CRC Press, 1st Edition, 2024
- Briggs & Kodnani, “The Potentially Large Effects of Artificial Intelligence on Economic Growth,” Goldman sachs, Economics Research, 26 March 2023.
- Brookings Institution, “How insurance can mitigate AI risks,” 2019. 11
- Gopal Mengi, “Accessible and Individualized Learning,” MIT Computer Science and Artificial Intelligence Laboratory (CSAIL) Cambridge, MA The ACM Magazine for Students, 29(3), 2023, 58-59 <<https://doi.org/10.1145/3589653>>

- Harrison Lee, Samrat Phatale , Hassan Mansoor, Thomas Mesnard , Johan Ferret, Kellie Lu, Colton Bishop, Ethan Hall, Victor Carbune, Abhinav Rastogi, Sushant Prakash, “RLAIF vs. RLHF: Scaling Reinforcement Learning from Human Feedback with AI Feedback,” arXiv:2309.00267v3 [cs.CL] 3 Sep, 2023
- Jeremy Jones, Robert D. Clark, Michael S. Lawless, David W. Miller & Marvin Waldman, “The AI-driven Drug Design (AIDD) platform: an interactive multi-parameter optimization system integrating molecular evolution with physiologically based pharmacokinetic simulations,” Journal of Computer-Aided Molecular Design, 38(14), 2024
- Kautz, H., “The third AI summer: AAAI Robert S. Englemore memorial lecture,” AI Magazine 43(1), 2022, 105–125.
- Keith D, Sean O, Kara T, Tanja B, Marcos B, “The \$100 Billion Opportunity for generative AI in P&C claims handling,” Bain & Company, 2024  
<<https://www.bain.com/isights/100-billion-dollar-opportunity-for-generative-ai-in-p-and-c-claims-handling>>
- Korinek, A., Balwit, A., , Aligned with Whom? Direct and Social Goals for AI Systems, NBER Working Paper Series, 2022, 30017.
- Loughran, T., McDonald, B., “When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks,” The Journal of Finance 66(1), 2011, 35–65.
- The Geneva Association (Noordhoek), “Regulation of Artificial Intelligence in Insurance: Balancing Consumer Protection Innovation” The Geneva Association, 2023  
<<https://www.genevaassociation.org/publication/public-policy-and-regulation/regulation-artificial-intelligence-insurance-balancing>>.
- Turing, A., “Computing machinery and intelligence,” Mind 59, 1950, 433–460.  
<<http://cogprints.soton.ac.uk/abs/>>

## Abstract

The development of generative AI is increasing, and its adoption in the insurance industry is on the rise. Accordingly, this study aimed to analyze which parts of loss assessment work are being replaced or supplemented by generative AI technology and how its role is changing.

As a result, the introduction of AI in the loss assessment process will first enable automated report generation, and with the automation of data analysis and report writing, loss assessors are likely to shift to the role of reviewing data and verifying AI results. Second, image-based accident assessment is possible, and this is mainly used in traffic accident damage assessment work. Third, through strengthened expertise, the scope of the role will expand and change to include complex claim processing, situational judgment in the event of ambiguity in terms and conditions, legal disputes, or differences in interpretation, negotiation and emotional management with insured persons, coordination of opinions with insurance companies, and fraud determination and investigation.

The introduction of AI brings with it both positive aspects and data, algorithmic, ethical, and social risks. AI risks are defined as the potential negative consequences that AI will have on individuals and businesses. If insurance companies assume these risks, their loss adjustment operations will require the following capabilities: Damage assessment work includes determining whether there was a possibility of predicting or avoiding AI errors, taking preventative measures and responding to errors, identifying and proving the source of negligence between AI developers, users, and external attackers, and calculating the amount of damages in the event of damage caused by AI errors.

In addition, it is necessary to determine whether AI-related damages are covered under the insurance policy, examine whether the accident was accidental or whether a software error is exempt from liability, and then

calculate the amount of direct and indirect damages. In conclusion, in AI risk, loss adjusters are required to go beyond simple compensation assessors and become integrated experts in technology, law, and insurance. This includes proving ethical and institutional legitimacy, making emotional and technical judgments, and even suggesting risk management and policy improvements.

※ **Key words** : Generative AI, AI in insurance companies, Changing role of loss adjusters, AI risks