KIRI 리 포 트
글로벌 이슈

딥페이크(Deepfake)와 보험사기

손민숙 연구원

요 약

딥페이크 기술의 발달로 정보의 진실성이 위협받고 있으며, 이는 향후 기업들의 새로운 위험요소로 평가되고 있음. 딥페이크는 영상, 이미지, 음성 파일 등 다양한 형태로 나타날 수 있는데 특히 보험산업은 보험금 청구 등에 사진과 같은 디지털 콘텐츠를 활용하고 있어 딥페이크 보험사기에 취약할 것으로 예상됨. 이에 적절한 대응을 위해 전체 보험산업 연합 차원의 기술 개발을 추진하는 것이 하나의 방안이 될 수 있으며 향후 기술적, 법적, 사회적으로 광범위한 혼합 조치 및 대응이 필요함

○ 최근 딥페이크(Deepfake) 기술이 발전하면서 정보의 진실성이 위협받고 있으며, 사이버 위험은 향후 기업이 마주할 수 있는 가장 심각하고 발생 가능성이 높은 새로운 위험요소 중 하나로 평가되고 있음[1)]

- 기업의 정보 보안에는 CIA(Confidentiality, Integrity, Availability)가 요구되는데, 최근 딥페이크 기술이 발전하면서 CIA 중 정보의 진실성(Integrity)이 위협을 받고 있음
- 딥페이크 동영상 개수는 2018년에서 2020년 사이 10배 이상 증가하였으며[2)], 사이버 위험 분석 전문회사인 CyberCube에 따르면 딥페이크 기술이 향후 2년 내 기업 등에 주요 위험요소가 될 것으로 전망됨[3)]
- 2019년 8월 CEO의 목소리를 흉내 낸 딥페이크 음성으로 인해 영국의 한 에너지 회사가 24만 3,000달러의 사기를 당했다는 최초의 뉴스 보도는 기업들에 경종을 울렸으며[4)], 이후 딥페이크에 대한 기업들의 관심이 높아짐

○ 딥페이크는 머신 러닝(machine-learning) 기술을 이용한 디지털 콘텐츠의 제작 또는 조작으로 정의될 수 있으며, 영상 또는 이미지, 텍스트, 음성 파일 등 다양한 형태로 나타날 수 있음[5)]

- 딥페이크는 기술적 요소를 기반으로 하며, 특히 머신 러닝을 이용한 딥페이크 생성 프로세스와 관련이 높음
- 디지털 콘텐츠의 조작은 새로운 현상이 아니지만 딥페이크의 생성 과정에 머신 러닝 기술을 적용함으로써 ① 콘텐츠 품질의 향상, ② 대규모 콘텐츠 제작, ③ 대중화가 가능해짐
- 딥페이크는 개인적·조직적·사회적으로 평판 및 금전적 손상을 초래하고, 의사결정 과정의 조작에 영향을 미치게 되므로 피해의 심각성, 피해 규모, 피해 대상의 탄력성을 고려한 적절한 대응이 요구됨

1) Marsh & McLennan(2021. 1), “DIGITAL DECEPTION: Is your business ready for “deepfakes”?”, MMC Cyber Handbook 2021
2) IRGC(2021. 5. 12), “Risk governance and the rise of deepfakes”, Spotlight on risk
3) CyberCube(2020), “Social Engineering, blurring reality and fake”
4) WSJ(2019. 8. 30), “Fraudsters Used AI to Mimic CEO’s Voice in Unusual Cybercrime Case”
5) IRGC(2019), “Forged Authenticity, Governing Deepfake Risks”

○ 특히 보험산업은 보험금 청구 프로세스에서 사진 등을 증거로 활용 중인데, 기존의 위험 평가모델 및 손해사정시스템으로는 딥페이크를 활용한 보험사기에 대응이 어려워 가짜 콘텐츠로 인한 문제가 심화될 것으로 예상됨[6)]

- 사이버 위험에 대한 우려로 사이버보험에 대한 관심 및 판매가 증가할 수 있으나, 눈앞에 보이는 '사실'이 더 이상 '진실'이 아니라는 것은 보험회사에도 동일함
- 딥페이크 사기로 과도한 금액이 청구되지 않는 한 일부 보험금 청구로 보험회사에 큰 타격이 있을 것이라고 볼 수는 없으나, 문제는 이로 인해 기업의 평판 또는 브랜드의 가치가 손상되는 경우임[7)]
- 상당한 수준의 법률 비용, 위기관리 비용, 손상된 시스템 복구 비용, 데이터 수정 비용 등이 발생할 것으로 예측되며, 특히 기업의 평판이 손상되는 경우 잠재적 수익의 감소 및 주가 하락이 초래될 것으로 예상됨[8)]
- 또한 딥페이크 사기로 인해 발생한 손실이 보험료에 반영되어 보험계약자의 부담을 가중시키고 불필요한 사회적 비용을 유발할 수 있음
- 따라서 이러한 가짜 콘텐츠를 식별하고, 고객 인증을 강화하며, 지속적인 보안 테스트를 실시하는 등 적절한 대응 체제를 갖추어야 하나, 투입 비용 대비 효용을 예측할 수 없어 시스템 마련이 어려운 상황임[9)]

○ 전체 보험산업 연합 차원에서 기술 개발을 추진하는 것이 하나의 방안이 될 수 있으며, 딥페이크 관련 법안을 마련하고, 사회적으로 교육 및 시스템을 구축하는 등 향후 기술적, 법적, 사회적으로 광범위한 혼합 조치 및 대응이 필요함

- 딥페이크로 인한 위험에 대응하기 위해 Adobe, Microsoft, BBC는 콘텐츠 출처 및 진위 입증을 위한 연합(Coalition for Content Provenance and Authenticity; C2PA)을 결성하여 콘텐츠의 출처를 안전하게 입증하고, 수정 이력을 확인할 수 있도록 하는 표준을 개발 중임[10)]
- 해외 주요국들은 딥페이크의 생성 과정 및 결과물에 투명성을 요구하고, 딥페이크를 범죄로 규정하거나 사적 소송권을 인정하는 등 관련 법안을 마련하는 중임
  - 미국에서 제안된 딥페이크 책임법(DEEP FAKES Accountability Act) 초안과 EU의 인공지능법(Artificial Intelligence Act)에는 딥페이크 콘텐츠 및 생성 시스템에 투명성을 요구하고 있음[11)]
  - 미국의 버지니아주는 성적 이미지 공유를 통한 성희롱 범죄에 딥페이크를 포함하였고, 캘리포니아주와 텍사스주는 공직 후보자를 대상으로 한 딥페이크의 제작 및 배포를 금지하였으며, 중국은 딥페이크 기술을 이용한 가짜 뉴스의 배포를 금지하는 규칙을 시행 중임[12)13)]
- 그 밖에도 미국은 딥페이크 탐지 기술의 연구를 지원하고 상업화를 장려하기 위해 딥페이크 관련 대회(Deepfakes Prize)를 개최하고 상을 부여하는 등 사회적 노력을 기울이고 있음[14)]

---

6) Attestiv(2021), "COMBAT DIGITAL PHOTO AND MEDIA FRAUD IN INSURANCE"
7) IRGC(2019), "Forged Authenticity, Governing Deepfake Risks"
8) Marsh & McLennan(2021. 1), "DIGITAL DECEPTION: Is your business ready for "deepfakes"?", MMC Cyber Handbook 2021
9) Marsh & McLennan(2021. 1), "PREPARE TO PROTECT YOUR CUSTOMERS' VOICES", MMC Cyber Handbook 2021
10) IRGC(2021. 5. 12), "Risk governance and the rise of deepfakes", Spotlight on risk
11) 딥페이크 자체를 금지하고 있지는 않으며, 표현의 자유와의 균형을 고려한 것으로 판단됨
12) IRGC(2019), "Forged Authenticity, Governing Deepfake Risks"
13) WilmerHale(2019. 12. 23), "First Federal Legislation on Deepfakes Signed Into Law"
14) WilmerHale(2019. 12. 23), "First Federal Legislation on Deepfakes Signed Into Law"