2023년 제9호 (통권4호)

데이터산업 동향 이슈 브리프

ISSUE BRIEF

2023.9

신뢰할 수 있는 AI를 위한 데이터 품질 요구사항

Kdata 한국데이터산업진흥원

데이터산업 동향 이슈 브리프

C/O/N/T/E/N/T/S

신뢰할 수 있는 AI를 위한 데이터 품질 요구사항

요약

- 본 데이터산업 동향 이슈브리프 2023년 제9호는 유럽위원회(European Commission)의 공동연구센터(JRC)의 컨퍼런스 및 워크숍 보고서 **「포용적이고 편향되지 않으며 신뢰할 수 있는 AI를 위한 데이터 품질 요구사항(DATA QUALITY REQUIREMENTS FOR INCLUSIVE, NON-BIASED AND TRUSTWORTHY AI)[1]」**를 요약·정리하였음
- **원문 보고서는 'Putting-Science-Into-Standards(PSIS, 과학을 표준에 담다)' 워크숍 내용을 근간으로 함.** PSIS 워크숍은 유럽표준화위원회(CEN), 유럽전기기술표준화위원회 CENELEC) 및 유럽위원회의 공동연구센터(JRC)가 공동 주최하며, 새로운 과학 및 기술분야의 표준화를 통한 혁신을 촉진하고 산업경쟁력을 강화하는 것을 목표로 함
- 2022년 6월 개최된 PSIS 워크숍은 포용적이고 편향되지 않으며 신뢰할 수 있는 AI를 위한 데이터 품질 요구사항에 초점을 맞춤
  - 표준화 활동의 우선순위, 특정 기술(Particular technologies), 잠재적 표준화 로드맵 초안 작성을 주 내용으로 진행함
- 이를 바탕으로 본 보고서는 첫째, **데이터세트의 생성 및 데이터화부터 데이터 품질 요구사항, AI 시스템의 편향성 검사 및 완화까지 기존 표준과는 다른 새로운 표준의 필요성을 강조함**
  - 새로운 표준 초안 작성을 위한 전체 프로세스를 확인하고, 새로운 표준은 업계, 중소기업 대표, 시민사회, 학계 등 모든 관련 이해관계자를 포용하고 충분히 대표하는 것이 중요함
  - AI법에 따라 AI 시스템의 시장 배포(Deployment) 및 성장 연구 분야 지원을 위한 AI 표준 개발에 전문가의 참여 확대는 필수적임
- 둘째, **6개의 특정 산업에서 AI 모델의 데이터 품질 요구사항과 잠재적인 표준화 요구사항에 대해 논의함**
  - 6개 산업은 소비자와 업계의 관심이 가장 높은 분야로 △ 교육 및 고용, △ 법 집행 및 공공 부문, △ 금융, △ 의료 및 헬스케어, △ 산업자동화 및 로봇, △ 소셜 미디어, 콘텐츠 중재, 추천 시스템을 포함한 미디어로 구성됨
- 본 보고서는 현재 진행 중인 논의 전개방향과, 중소기업의 관점, 정책적 관점, 입법자, 소비자 보호, 기본권, 기초과학적 관점 등에서 산업별 정보를 결합에 대해 향후 논의가 필요함을 제시함

1) Joint Research Centre (European Commission), 「DATA QUALITY REQUIREMENTS FOR INCLUSIVE, NON-BIASED AND TRUSTWORTHY AI」, (2022) (https://op.europa.eu/en/publication-detail/-/publication/b11a0504-75eb-11ed-9887-01aa75ed71a1/language-en/format-PDF/source-291783199)

PART Ⅰ 개요

- 인공지능(AI) 분야는 지난 10년간 전례 없는 속도로 발전해 왔으며, 의료, 금융, 제조, 소셜미디어, 법 집행, 환경, 농업 등 다양한 분야에 걸쳐 폭넓게 실용화됨. 데이터는 입력된 학습 데이터에 맞춰 출력하는 적응형('학습') 데이터 처리 알고리즘인 AI 시스템의 근원적인 역할을 함. 이러한 데이터의 역할은 유럽연합(EU) 정책 의제에도 반영됨

  - AI 및 데이터에 대한 지속적인 정책 개발의 예는 다음과 같음

    - 유럽연합은 유럽 데이터 전략(European Data Strategy), 유럽 일반 개인정보보호법(GDPR, General Data Protection Regulation), 디지털 서비스법(Digital Services Act), 데이터 거버넌스법(Data Governance Act), AI법(AI Act), 데이터법(Data Act) 등이 있음

  - 유럽위원회 공동연구센터(JRC, Joint Research Center)는 2022년 6월 8일, 유럽표준화위원회(European Committee for Standardization) 및 유럽전기기술표준화위원회(European Committee for Electrotechnical Stadardization)와 공동으로 2022 과학표준화워크숍(PSIS, Putting Science Into Standards workshop)을 개최함

    - AI 시스템 개발에서 데이터의 중요한 역할과 AI용 데이터 표준화를 향한 여정을 고려하여, 2022 PSIS에서는 '포용적이고 편향되지 않으며 신뢰할 수 있는 AI를 위한 데이터 품질 요구사항'이라는 주제에 초점을 맞춤

    - 해당 워크숍에는 36개국 178명 이상의 유럽 표준화 전문가, 정책 입안자, 과학자 및 여러 이해관계자가 참석했으며, 워크숍에 참가한 조직은 민간 부문 36%, 공공 부문 23%, 표준화 기구 19%, 학계 18%, NGO/노동조합/시민사회단체가 4%로 다양한 유형의 기관이 분포됨

    - 워크숍의 주요 목표는 다음과 같음
      i) AI 및 향후 AI법의 맥락에서 데이터 품질 및 관련 윤리적 문제를 해결하기 위한 현재와 미래의 요구사항과 권장 사항 모색
      ii) 표준화 진행과 AI 시스템 개발자에게 장애물이 되거나 AI 시스템의 신뢰성에 영향을 미칠 수 있는 잠재적인 표준화 격차 식별
      iii) AI 모델에 대한 데이터 품질 및 관련 표준을 보장하기 위한 핵심요소 수집
      iv) 표준 또는 기타 관련 문서(특정 산업의 요구사항 해결을 위한 지침 문서 또는 과학적 측면을 설명하는 기술 보고서 등) 개발의 우선순위와 일정 파악

- 그러나 AI의 이점은 디지털 편견에 의한 차별과 불평등을 초래할 수 있는 위험성을 내포함. 이에 따라, 신뢰할 수 있는 AI 개발을 보장하기 위해 데이터 품질 표준을 제정할 필요성이 대두됨

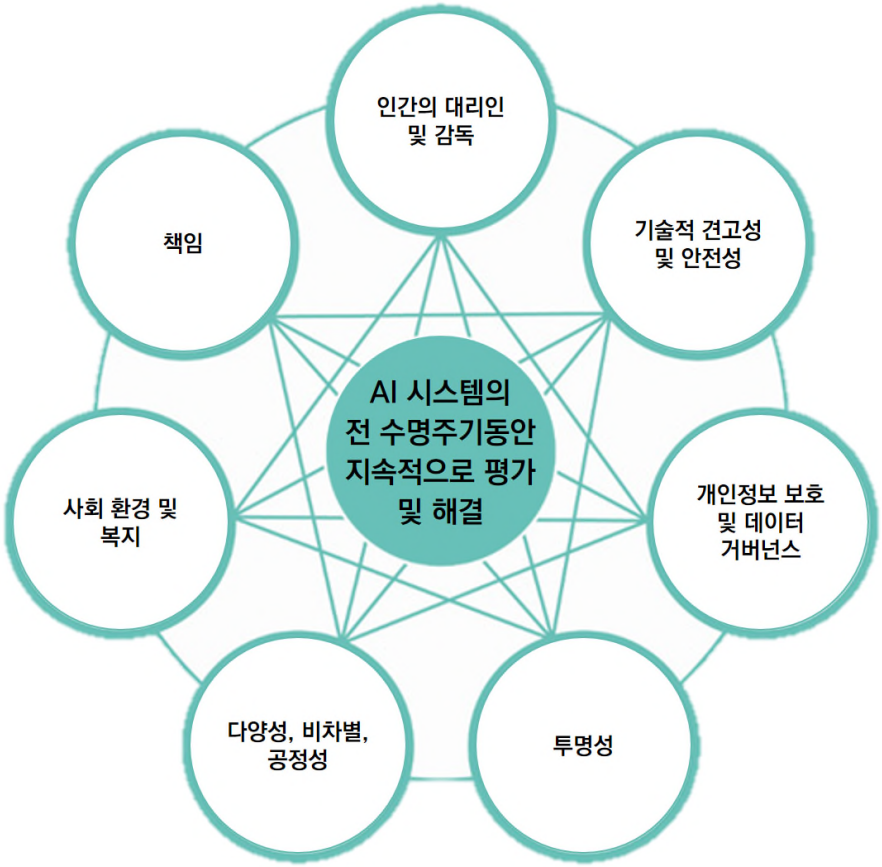
  - 2018년 AI 고위급 전문가 그룹(AI HLEG, AI high level expert group)은 신뢰할 수 있는

AI 시스템을 위한 7가지 핵심 요건 중 하나로 '다양성(Diversity), 비차별(Non-discrimination), 공정성(Fairness)'을 명시함

- 이 요건은 불공정한 편견을 제거하고, 누구나 AI 제품 및 서비스를 사용하도록 접근성과 유니버설 디자인 원칙을 고려해야 하며, 프로세스 전반에 걸쳐 영향을 받는 모든 이해관계자와의 협업을 강조함

- 이외에도 기술적 견고성과 안전성, 개인정보 보호 및 데이터 거버넌스, 책임성과 같은 핵심 요건도 데이터 품질과 연관이 있음

| 그림 1 | 신뢰할 수 있는 AI 시스템의 7가지 요구사항의 상호관계[2)]



2) European Commission, Requirements of Trustworthy AI, https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines/1.html

PART Ⅱ 국제 권고사항 및 표준화 동향

- 유네스코(UNESCO)가 2021년에 채택한 '인공지능의 윤리에 관한 권고안'은 회원국이 다양성과 포용성 증진을 위한 정책을 시행하도록 촉구하며, AI 기술과 학습 데이터세트 개발에 자국 인력을 투입할 것을 권고함
  - 해당 권고안은 개방적이고 신뢰할 수 있으며, 최적의 표준 데이터세트(Gold Standard Dataset) 생성의 투자를 포함하며, 필요한 경우 데이터 주체의 동의 등 유효한 법적 근거를 바탕으로 구성하도록 함
    - 또한, 권고안의 이행을 촉구하기 위한 역량 강화 및 평가 메커니즘, 특히 준비성 및 윤리적 영향 평가 도구를 포함하고 있음. 이 도구는 AI 수명주기 전반에 걸쳐 적용되며, 개발 및 배포와 관련된 모든 주체들이 참여하여 AI 기술 정책을 투명하게 모니터링하고 평가할 수 있도록 함
- OECD AI 원칙(2019)은 신뢰할 수 있는 AI를 위한 책임있는 관리 원칙과 국가 정책 및 국제 협력이라는 두 가지 영역과 관련된 10가지 원칙을 수립함
  - 이 원칙의 이행을 위해 진행 중인 프로젝트는 다음과 같음
    - OECD AI 시스템 분류 프레임워크: OECD 원칙이 공공 정책 영역에 미치는 영향에 따라 AI 시스템을 분류하는 데 도움이 되는 프레임워크임
    - 전 세계에서 발생하는 AI 사고 및 논란 정보 수집을 위한 AI 사고 추적: 효과적인 AI 정책을 위한 증거 기반을 만들고, 사건의 재발을 방지하고자 함
    - 신뢰할 수 있는 AI를 위한 도구 카탈로그: 사람들이 AI 시스템이 OECD 원칙을 준수하는지 확인하는 데 사용할 수 있는 도구와 체계적인 방법을 제시함
    - AI 책임성 및 위험평가: AI의 책임성의 주요 구성 요소의 일반적인 이해와 기존 이니셔티브(예: ISO, IEEE, NIST, CEN-CENELEC)에 기여하고, OECD의 AI 작업과 위험 관리 프레임워크를 통합하는 것을 목표로 함
- 유럽연합의 AI법(AI Act)은 허용하는 AI 기술과 이러한 기술이 유럽연합 시장에 배포될 수 있는 조건을 정의하고자 함. 이 목표는 위험 기반 접근 방식을 통해 달성할 수 있으므로 AI법은 주로 고위험 AI 시스템에 초점이 맞춰짐
  - 일관되고 높은 수준의 안전 및 기본권 보호를 위해 AI법은 모든 고위험 AI시스템에 대한 필수 요건을 제안함. 제안된 요건의 준수 여부는 적합성 평가 절차를 통해 입증되어야 하며, 특정

상황에서는 제3자 인증 기관의 개입이 필요할 수 있음

- AI법은 제품법 규제 체계에 따라 구성되어 기술적 목표 및 요건만 주요 법률에 포함하고, 이러한 요건의 준수를 입증하기 위한 보다 세분화되고 구체적인 기술 솔루션은 주로 유럽 표준화 기구에서 개발한 조화된 표준(Harmonised Standards)[3]을 통해 설정될 것으로 예측됨

- 따라서, 데이터 품질과 관련하여 AI법이 최신 기술을 반영하는 효과적인 기술 솔루션으로 발전하기 위해서는 조화된 표준의 역할이 핵심임

- 특히 데이터 표준화와 관련하여, 유럽위원회는 표준화 프로세스에 관련 이해관계자의 적절한 참여를 보장하는 것이 중요함을 강조함. 예를 들어, 중소기업이 중요한 역할을 하는 AI 기술에서는 대기업과 중소기업의 참여가 필수적이며, 의료 AI의 데이터 품질과 관련된 표준 개발에는 의료 전문가와 환자의 견해가 중요함

- 또한, AI법은 시행 이후의 상황도 중요하며 이는 시장 출시 후 공급자의 모니터링으로 구체화됨. 데이터와 관련하여 시장 출시 전후 통제를 효과적으로 수행하기 위해 AI법은 인증기관과 국가 당국에 공급자가 활용하는 데이터세트의 접근 권한을 명시적으로 부여함

국제 표준화 동향

- 국제 AI 표준화 측면에서 ISO/IEC JTC1 SC42(인공지능 표준화 회의)는 AI 작업의 기초 표준(예: 용어 등)부터 AI 관리 및 신뢰성에 관한 표준에 이르기까지 다양한 단계에서 상당한 수의 표준을 발표하거나 개발 중임

  - 구체적인 표준으로 용어, 데이터 품질 측정, 관리 요구사항 및 지침, 품질 프로세스 프레임워크 등 데이터 품질을 포괄적으로 다루는 5259 시리즈(ISO 2022)의 표준이 있음

  - 윤리 및 사회적 문제를 다루는 TR 24368 초안이나 AI 시스템의 편견에 관한 TR 24027초안(2021) 등이 있음(ISO/IEC 2022)

- 전기전자기술자협회(IEEE)도 윤리적 측면에 중점을 둔 7000시리즈를 포함하여 광범위한 AI 표준을 개발하고 있음

  - 이 시리즈 중 특히, 알고리즘의 편향성과 데이터품질을 다루는 IEEE 7003 표준은 AI 제공업체가 설계 프로세스에 윤리적 고려사항을 통합할 수 있도록 지원하는 방법론부터 투명성, 편향성, 개인정보 보호 관련 표준까지 그 범위가 넓으며, 인증 프로그램에 대한 구체적인 작업도 포함되어 있음

- 미국 국립표준기술연구소(NIST, National Institute of Standards and Technology)도 AI의 편향성을 관리하고 AI 위험 관리 프레임워크를 개발하는 작업을 진행함

3) 조화 표준은 공인된 유럽 표준 기구(CEN, CENELEC 또는 ETSI)에서 개발한 유럽 표준으로 제조업체, 기타 경제 운영자 또는 적합성 평가 기관은 조화 표준을 사용하여 제품, 서비스 또는 프로세스가 관련 EU 법률 준수 여부를 입증함 (https://single-market-economy.ec.europa.eu/single-market/european-standards/harmonised-standards_en)

- NIST는 가용성만을 기준으로 데이터를 선택하거나 데이터 시스템을 최적화된 시나리오로 제한하는 등 잠재적인 편향의 원인을 인식하면서 맥락을 고려하여 피해를 식별, 평가 및 예방하는 위험 관리 프레임워크를 개발 중임

● 유럽 표준화 기구인 CEN과 CENELEC는 2021년 AI 표준 개발 및 채택을 담당하고 AI와 관련된 다른 기술 위원회에 지침을 제공하는 역할을 하는 공동 기술 위원회 21 '인공지능'(JTC 21)을 출범시킴

- JTC 21은 유럽 표준과 국제 표준의 조합이 필요한 경우, 두 표준 사이의 격차를 유럽 수준에서 해결하는 역할을 함

- 보완적인 유럽 표준화 범위는 AI 신뢰성 특성(인적 감독, 정확성, 견고성 등), 데이터 품질, 위험 관리에 대한 구체적인 지표 및 통제 제공, 고위험 AI 시스템의 적합성 평가 방법론 등을 포함할 것으로 예상함

- 이를 통해 유럽 AI 규정의 수평적 특성에 따라 전반적으로 고위험 사례에 적용가능한 표준 기반이 마련될 것이며, 서로 다른 사양 간의 완전한 일관성을 보장하게 될 것으로 기대함

| 표 1 | 국제 권고사항 및 표준화 동향 간략 요약

	국제 권고사항 및 관련 법안			국제 표준화 동향
	유네스코(UNESCO) AI 윤리 권고안(2021)	OECD AI 원칙(2019)	유럽연합의 AI법 (AI Act)	
기본 방향	- 회원국의 다양성과 포용성 증진을 위한 정책 시행 촉구 - AI 기술과 학습 데이터세트 개발에 자국 인력 투입 권고	- AI를 위한 책임있는 관리 원칙과 ,국가 정책 및 국제 협력이라는 두 가지 영역과 관련된 10가지 원칙 수립	- 허용되는 AI 기술과 해당 기술이 유럽연합 시장에 배포될 수 있는 조건을 정의 - 주로 고위험 AI 시스템에 초점	- AI 작업의 기초 표준(예: 용어 등)부터 AI 관리 및 신뢰성에 관한 표준에 이르기까지 다양한 단계에서 상당한 수의 표준 발표 및 개발
세부 요건	- 개방적이고 신뢰할 수 있으며, 최적의 표준 데이터세트(Gold Standard Dataset) 생성에 투자 - 필요 시 데이터 주체의 동의 등 법적 근거가 바탕이 되어야 함	- OECD AI 시스템 분류 프레임워크 구동 - AI 사고 추적 시스템 구동 - 신뢰할 수 있는 AI를 위한 도구 카탈로그 사용 - AI 책임성 및 위험평가 시행	- 모든 고위험 AI 시스템에 대한 필수 요건 제안 - 요건 준수 여부는 적합성 평가 절차를 통해 입증 - 필요 시 제3자 인증 기관의 개입	-
관련 사례	-	-	-	- ISO/IEC JTC1 SC42(인공지능 표준화 회의) - 5259시리즈(ISO 2022) 표준 - 전기전자기술자협회(IEEE)/ 7000시리즈 등

PART III 연구 및 산업 분야의 데이터 품질 평가 및 편향성 완화를 위한 수평적 이니셔티브

- 본 워크숍의 첫 번째 파트에서는 업계와 학계의 수평적 데이터 품질 및 투명성의 접근 방식에 초점을 맞춰 데이터세트 구축, 모델 학습, 시스템 배포, 시장 출시 후 모니터링 등 AI 시스템 수명주기의 여러 단계를 다룸
  - 수평적 이니셔티브는 성별, 국적, 문화, 언어, 학문 분야를 고려한 다양한 그룹의 데이터 편향성을 포괄함. 첫 번째 파트에서 다루는 두 가지 수평적 접근 방식은 1) AI를 위한 데이터세트의 생성 및 문서화(Documentation), 2) AI 데이터 품질과 편향성 검사(Examination) 및 완화(Mitigation)의 측면임

2.1 AI용 데이터세트 생성 및 문서화

- 데이터는 AI 시스템을 학습, 테스트 및 검증하는데 필요한 원재료로 데이터 준비, 큐레이션, 주석 달기, 정리, 공유 등 AI 개발 수명주기 전반에 걸쳐 핵심적인 역할을 함
  - 완전성, 정확성, 대표성, 개인정보 보호의 측면에서 고품질 데이터세트의 생성은 신뢰할 수 있는 AI 시스템의 개발에 직접적인 영향을 미침
- 그러나 최근 연구에 따르면, AI에 널리 사용되는 대부분의 데이터세트는 노이즈(레이블이 잘못 지정된 샘플, 누락된 데이터, 저품질 샘플 등)가 많고, 편향되어 있으며(특정 소수 집단을 거의 포함하지 않는 인구통계학적 측면), 데이터세트의 품질이 낮은 것으로 밝혀짐
  - 학습 데이터세트에서 노이즈 수준이 증가할수록 AI 시스템의 성능이 명백하게 저하되는 것이 확인됨
    - 예를 들어, 라벨이 잘못 지정된 샘플을 10% 수동으로 수정하면 데이터세트의 크기를 두 배로 늘린 것과 비슷한 수준의 성능 향상을 가져온다는 사실이 입증됨
  - 이에 따라, 신뢰할 수 있는 데이터 세트의 생성 및 문서화를 위한 이니셔티브가 추진되고 있음
    - '데이터세트를 위한 데이터시트(Datasheets for Datasets)[4]'를 시작으로 업계와 학계의 다양한 이해관계자들이 데이터세트의 생성 과정과 내용의 투명성을 높이기 위해 여러 유형의 문서화 접근법을 제안함
    - 대표적인 문서화로는 설문지(Questionnaires: Gebru, et al. 2018, OECD 2022),

4) Gebru, T, J Morgenstern, B Vecchione, J Wortman Vaughan, H Wallach, H Daume III, and K Crawford., 「Datasheets for Datasets」(https://arxiv.org/abs/1803.09010), (2018)

체크리스트(Checklists: Madaio, et al. 2020, European Commission 2019), 정보 시트 또는 카드 (Information sheets/cards: M. Mitchell, et al. 2019, Matthew, et al. 2019), 위젯(Widgets: Chmielinski, et al. 2022, Bäuerle, et al. 2022, Holland, et al. 2018)의 형태가 제시됨

2.1.1 최신 기술, 도전 과제 및 지속적인 표준화 활동

AI 시스템, 특히 기본 AI 모델이 딥러닝을 기반으로 하는 경우, 방대한 양의 데이터가 필요함. 이에 데이터세트의 수집, 생성 및 문서화와 관련하여 시급히 해결해야 할 6가지 주요 과제를 논의함

1) 데이터 수집과 준비 과정은 시간이 많이 소요되고 오류가 발생하기 쉬움

- 데이터 과학자는 AI 시스템의 학습 및 검증을 위해 데이터를 준비하고 라벨링하는 데 약 60%의 시간을 소요하며, 데이터 수집 과정에 두 번째로 많은 19%의 시간을 소요함
- 따라서 데이터 수집 및 준비는 AI 시스템 개발 파이프라인의 약 80%를 차지하고, 이러한 작업은 여러 사람의 감독과 수작업으로 진행되므로 오류가 발생할 가능성이 높음

2) 웹스크래핑(Web scraping)은 잠재적으로 위험을 수반하는 관행임

- 웹스크래핑은 봇(Bot)을 사용하여 웹사이트에서 데이터를 자동으로 추출하는 프로세스로, AI를 위해 대량의 데이터를 수집하는 가장 인기있는 방법 중 하나임
- 이러한 관행은 저작권 문제와 데이터세트의 출처를 추적하는 데 어려움을 수반하게 되며, 특히 신뢰할 수 없는 출처에서 데이터를 수집하면 악의적인 고정관념과 검색 엔진의 편견이 지속되어 편향성이 증대될 수 있음

3) 데이터세트는 실제 사용 사례를 대표하지 못함

- 현재 AI용 데이터세트는 AI 시스템이 배포되는 실제 상황을 대표하지 못함
- 대표적으로 인구통계학적 편향(예: 인구통계학적 데이터가 소수 집단을 배제하고 백인 집단만 대표함), 고려되지 않은 코너 케이스(예: 데이터세트에 소수 그룹이나 계층이 거의 대표되지 않음), 데이터세트와 운영 데이터 간의 품질 격차가 있을 수 있음

4) 데이터세트 평가 지표가 일반적이고 연구지향적임

- 데이터세트 품질은 샘플 수, 데이터 상관관계, 누락된 데이터 비율 및 레이블이 잘못 지정된 샘플 비율과 같은 일반적인 기준을 통해 간단하게 평가함
- 데이터의 다양성, 대표성, 편향성 또는 데이터 콘텐츠의 윤리적 위반과 같은 요소를 포함하는 새로운 AI 전용 품질의 지표를 고려해야 함

5) 잘 정의된 데이터 거버넌스, 큐레이션[5] 및 문서화 프로세스가 부족함

- 데이터 세트의 품질과 무결성은 생성 과정뿐만 아니라 전 수명주기 동안 보장되어야 함

5) 디지털 자산을 선택, 보존, 유지, 수집, 아카이빙하는 작업

- 비정형적이고 복합적인 대규모 데이터세트를 문서화하는 과정은 매우 어렵고 모든 수명주기 단계를 다루기에는 한계가 있으므로, 도메인에 특화된 보다 수직적인 접근 방식을 도입해야 함

6) 문서화 관행과 데이터 공유를 촉진하기 위한 인센티브가 부족함

- 많은 기업에서 데이터를 투명하게 공유하는 것이 자사의 이익에 반한다고 인식하여 데이터 문서화 및 공유가 더딤. 특히 대기업은 대규모 데이터세트의 통제력을 보유하여 중소기업 간의 AI 발전 격차를 더욱 심화시킴

- 공공 부문에서도 공유할 수 있는 대량의 데이터를 보유하고 있으나, 민간과 공공 부문 사이에서 제대로된 데이터 공유 메커니즘을 구축하고 있지 않음

이러한 문제를 해결하기 위해 이미 진행 중인 활동과 이니셔티브를 바탕으로 한 해결책이 필요함

● 데이터세트 및 AI 문서화 방법론

- AI용 데이터에 초점을 맞춘 '데이터세트를 위한 데이터시트(Datasheets for Datasets)'와 '데이터 뉴트리션 레이블(Data Nutrition Label)[6]'은 전 세계 AI 실무자들이 폭넓게 채택하고 있는 구조화된 접근법임

- 그 외에도 모델 카드(Model Cards), AI 팩트시트(AI Factsheets), OECD 프레임워크(OECD Framework)와 같은 방법론도 데이터 품질 관련 고려사항을 제시하며, 향후 표준을 위한 좋은 기반이 될 수 있음

● 프로그래밍 툴킷

- 문서를 자동으로 생성할 수 있는 여러 프로그래밍 툴킷을 도입하면 데이터세트 문서화의 작업량을 줄일 수 있음

- 대표적으로 심포니(Symphony)[7], 모델 카드 툴킷(Model Cards Toolkit)[8], 데이터 뉴트리션 프로젝트 (Data Nutrition Project)[9] 등이 있으며, 대부분 데이터 중복, 이상, 편향성, 잘못 레이블링된 데이터 해결 등 문제를 해결할 수 있음

● 체크리스트

- 체크리스트는 AI시스템의 수명주기 동안 관련 모든 데이터의 품질 요소가 충족되었는지 확인하기 위해 감독관뿐만 아니라 AI 개발자 및 제공자에게도 유용한 형식임

- 예를 들어, 유럽위원회의 AI 관련 고위급 전문가 그룹이 작성한 '자체 평가를 위한 신뢰할 수 있는 AI 평가 목록'에서 제안하는 AI 공정성 관련 체크리스트[10]가 있음

6) 데이터 뉴트리션 레이블(Data Nutrition Label)은 데이터 사용자에게 제공되는 데이터의 품질, 출처, 정확성, 개인정보 보호 및 기타 중요한 정보를 간결하게 표시하는 방법을 설명하는 개념. 이 아이디어는 식품 상품에 표시되는 영양 정보 레이블에서 영감을 받았으며, 데이터의 뉴트리션 레이블을 사용자가 더 잘 이해하고 데이터의 신뢰성과 유용성을 평가할 수 있도록 돕기 위해 도입됨

7) https://www.whysymphony.com/

8) https://www.tensorflow.org/responsible_ai/model_card_toolkit/guide

9) https://datanutrition.org/

● **신뢰할 수 있는 AI를 위한 업계 이니셔티브**

- 신뢰할 수 있는 AI 시스템의 개발 및 상용화를 촉진하기 위해 현재 진행 중인 대표적인 이니셔티브로는 독일 전기기술자협회(VDE)의 '인공지능 신뢰 라벨(AI Trust Label)[11]'이 있으며, 이는 AI 제품에 신뢰 라벨을 부착하기 위한 사양을 정의함

- '프랑스 컨피안스 AI(Confiance.ai)[12]'는 신뢰할 수 있는 데이터 세트를 얻기 위한 방법, 도구 및 협업 메커니즘 개발에 중점을 둠

2.1.2 기존 표준화 연구의 공백과 표준화 기회

▶ 현재 표준이 AI 시스템을 학습하고 검증하는데 사용되는 데이터세트와 관련한 모든 문제를 해결하지 못하므로 'AI용 데이터'에 대한 구체적인 표준의 필요성이 대두됨

● 표준은 대표성을 고려해야 하며, AI 시스템이 배포될 운영 조건과 인구통계학적 환경을 충분히 대표하는 데이터를 확보해야 함

- 표준은 AI 내에서 데이터의 수집, 저장, 라벨링, 접근, 공유 및 사용 방법에 대한 방법론 개발을 통해 기존의 미흡한 프로세스를 발견하고 대처할 수 있어야 함. 이에는 표준화된 접근 가능한 데이터세트의 문서화 지침의 정의도 포함됨

● 데이터의 품질을 효율적으로 측정하기 위해서는 가이드라인도 필요함

- 사회적, 윤리적, 다양성 차원을 넘어 새로운 차원을 고려하기 위해 새로운 AI 전용 데이터의 측정 기준을 도입해야 함

- 원시 데이터에서 측정 기준을 자동으로 계산하고 모델 오류와 데이터 품질 문제의 상관관계를 식별하는 표준화된 툴킷은 AI 실무자가 데이터로 인한 문제를 식별하는 데 유용한 도구가 될 수 있음

● 알고리즘 표준화 도구는 정량적인 지표의 계산에 국한되어서는 안 되며, 데이터 큐레이션, 라벨링 및 문서화와 같은 정성적인 작업에도 도움이 될 수 있음

- 프로세스, 측정 기준 및 분류 체계 측면에서도 유럽 AI법과 같은 규제를 준수한다면 관련 체크리스트 뿐만 아니라, 특히 중소기업이 법적 요건을 준수하는 과정에서 도움을 받을 수 있음

▶ AI의 영역에 따라 다른 데이터 표준이 필요하며, 용어(Terminology), 측정(Metrology), 성능 특성화(Performance characterization), 호환성(Compatibility) 및 규제 평가(Regulatory assessment) 총 5가지 사항에 따른 표준화 요구사항은 다음과 같음

10) https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment
11) https://www.vde.com/de/presse/pressemitteilungen/ai-trust-label
12) https://www.confiance.ai/

● 1) 데이터세트 품질 평가를 위한 표준 체크리스트 정의, 2) 데이터 계통 매핑을 위한 방법론 구현, 3) 통합 분류 체계 및 데이터 라벨링 체계의 정의, 4) 데이터 문서화, 라벨링, 정리 및 메트릭 추출을 위한 표준화된 툴킷 구현이 우선순위가 높은 표준화 기회로 선정됨

- 1, 2번 항목은 기존의 사전 표준화 이니셔티브를 적극 활용할 수 있고 비교적 적은 노력으로 구현할 수 있음. 반면 3, 4번 항목은 아직 합의가 부족하고 배포되는 영역에 따라 데이터 유형이 매우 다양하며, 이를 운영하기 위해서는 강력한 구현 노력이 필요함

- 표준의 정립은 양질의 데이터를 보장하기 위해 중요하지만, 공동체가 이에 대한 필요성에 공감하고 데이터 품질과 데이터 공유 가치 및 문화를 창출하는 것이 더욱 중요함

| 표 2 | 데이터세트 생성 및 문서화를 위한 선별된 표준화 범주와 AI 가치사슬에 따른 사전 규범 연구 및 표준 요구사항 개요

	용어	측정	성능 특성화	호환성	규제 평가
AI 시스템 배포과 마케팅 - 규제 평가 - 사용자 - 투명성 및 명세서 - 책임 및 의무 - 유지 보수, 시장 후속 조치 및 편향성 모니터링 - 공급망					
AI 시스템 생성 및 생산 - 데이터 세트 및 알고리즘 (편향성 포함)/모델 - 사이버보안 - 시스템 설계 및 통합 - 업스케일링 및 평가 - 품질 관리		- 자동화된 도구 (데이터+모델 행동 분석) - 새로운 차원 (윤리적, 사회적, 다양성)			
데이터 생성 - 컴파일링, 준비, 편향성 검사 - 분석, 가공, 라벨링 - 라이선스 및 제한사항 - 공유 및 마케팅	- 통합분류체계 - 데이터 생성 과정의 통합된 부분으로서의 문서화 - 데이터 생성 및 수집에 관한 표준화된 과정 - 데이터 문서화 및 생성을 위한 자동화 툴	- AI 관련 데이터 품질 측정 기준에 대한 표준 - 공정성 및 편향성 지표	- 수명주기별 데이터 무결성 표준 - 대표성 보장 (운영 설정, 실제 데이터, 인구통계학적 설정)	- 데이터 형식 및 라벨링에 관한 표준 - 데이터 공유에 관한 표준	- 감사를 위한 데이터 품질 체크리스트 - 데이터 계통 증명 매핑 - 데이터 유지 관리

2.2 AI의 데이터 품질 및 편향성 검사 및 완화

- 데이터 품질과 편향성 검사 및 완화에 중점을 둔 AI 분야의 방법론, 도구 및 모범 사례에 대한 최신 연구를 발표함

  - 범위는 데이터 증강(Data augmentation), 가중치 손실 함수(Weighting loss functions, 예:인구통계 기반), 블라인드 방식(Blinding methods, 예:보호변수), 피드백 루프(Feedack loops), 미세 조정(Fine-tuning), 연합 학습(Federated learning), 전이 학습(Transfer learning), 견고성(Robustness), 평가 기법 및 벤치마크(Evaluation techniques and benchmarks), 적대적 공격(Adversarial attacks), 설명 가능성(Explainability), 인적 감독(Human oversight), 시판 후 모니터링(Post-market monitoring) 등이 포함됨

  - 데이터 품질 측정을 제대로 하더라도 AI 시스템 수명주기의 학습 및 배포 단계에서 비의도적인 편향성이 나타날 수 있으며, 이로 인해 예기치 못한 유해한 결과가 발생할 수 있음. 따라서, 알고리즘 학습, 배포 및 운영 기간 지속적인 편향성 검사 및 완화를 위해 모범 사례를 참고할 필요가 있음

  - 학습 단계에서 데이터 조작 및 증강 기술, 미세 조정, 특정 알고리즘 기법의 사용을 통해 편향성이 AI 시스템에 유입될 수 있으며, 배포 단계에서도 잠재적 피드백 루프, 부적절한 평가 기법 및 벤치마크, 또는 적대적 공격으로 인해 편향성에 취약함

  - 전체 AI 시스템 수명주기에서 편향성 및 기타 데이터 품질 문제를 감지하고 해결하려면 공정성, 투명성, 설명 가능성, 견고성 및 지속 가능한 모니터링 및 평가 기법에 관한 구체적 모범 사례를 채택해야 함

2.2.1 기존 표준화 연구의 공백과 표준화 기회

- AI의 중요한 애플리케이션을 정의하기 위해서는 데이터 품질에 관련된 법적 프레임워크에서의 공정성이 중요함

  - 보증 사례(Assurance cases)[13]와 그 식별은 신뢰할 수 있는 데이터를 위한 매우 중요한 소스임. 규범기반 규정에서 목표기반 규정으로 전환하고 보증 사례를 기반으로 표준화하는 것이 이상적임

  - 공정성 문제 해결을 위해 데이터 중심 관점과 모델(알고리즘 코드) 중심 관점 사이의 대조를 핵심요소로 다루어야 함

    - 1) AI 편향을 막기 위해서는 스마트 데이터를 얻기 위한 데이터 처리, 2) 데이터 공정성을 통한 모델의 편향성 방지, 3) 데이터 프라이버시 보호가 주요 과제로 강조됨. 이를 위해 AI 감독 기관 설립, 데이터 품질 보장을 위한 방법론 개발, AI 사용자의 데이터 중심 접근을 위한 데이터 사일로 생성 등을 해결해야 함

13) 정보시스템이 특정 품질 요구사항을 충족한다는 것을 보여주는 구조화된 주장 또는 증거를 의미함(출처: csrc.nist.gov)

- AI 가치사슬에서 데이터 품질 표준, 편향성 식별 및 완화를 위해 적절한 데이터 품질 평가 지표에 합의가 필요하며, AI 시스템의 공정성과 지표에 대해 정의할 필요가 있음
  - 그러나, 데이터의 적합성을 측정하기 위한 AI 모델의 편향성 감지 및 측정에 대한 정의가 부족함. 더 많은 시스템 통합을 위해 복잡한 상황에서 AI 시스템의 상호 운용성을 허용하는 공통 언어, 사용 중인 교차 도메인에 대한 합의가 필요함
  - 또한, AI 시스템 배포의 전제 조건으로 목표 기반 규제 평가와 투명성을 문서화하는 표준을 개발해야 함. 복잡한 시스템에서 윤리적 영향을 평가하고 잠재적 편향과 보호 변수를 감지하기 위해 사전 규범 연구가 필요함

| 표 3 | 데이터 품질과 편향성 식별 및 완화를 위한 선별된 표준화 범주와 AI 가치사슬에 따른 사전 규범 연구 및 표준 요구사항 개요

	용어	측정	성능 특성화	호환성	규제 평가
AI 시스템 배포과 마케팅 - 규제 평가 - 사용자 - 투명성 및 명세서 - 책임 및 의무 - 유지 보수, 시장 후속 조치 및 편향성 모니터링 - 공급망				- 도메인 간 사용을 위해 배포된 경우를 포함하여 다양한 상황에서 AI 시스템을 사용하기 위한 표준 정의	- 목표 기반 규제 평가 정의 - 배포 조건으로 문서화 표준 및 시스템 투명성 확보
AI 시스템 생성 및 생산 - 데이터 세트 및 알고리즘 (편향성 포함)/모델 - 사이버보안 - 시스템 설계 및 통합 - 업스케일링 및 평가 - 품질 관리	- AI 시스템의 공정성 정의	- AI 모델의 편향성 감지 및 측정 방법 정의 - AI 시스템 상황에 맞는 데이터 적합성 측정 - AI 시스템의 공정성 측정 지표 정의	- AI 시스템의 수명주기 전반에 걸친 보증 사례 - 복잡한 시스템에서 사용되는 AI의 윤리적 영향을 평가하기 위한 사례 연구 수행 - 시스템 편향성 테스트를 위한 벤치마킹 데이터세트 - 숨겨진 편향과 보호변수의 테스트		
데이터 생성 - 컴파일링, 준비, 편향성 검사 - 분석, 가공, 라벨링 - 라이선스 및 제한사항 - 공유 및 마케팅		- 데이터 품질 평가 지표			

PART Ⅳ 6대 산업 분야의 데이터 품질 요구사항

- (필요성) AI법은 디지털 단일 시장의 파편화를 방지하고, 다양한 산업분야의 AI 관련 규정이 조화롭게 유지될 수 있는 수평적 프레임워크를 제시함
  - 일반적으로 AI시스템의 개발주체(공공 또는 민간) 또는 배포된 산업분야와 무관하게 동일한 기술은 기본적인 권리(자율성, 데이터 종속성, 불투명성 등)에 대한 공통된 문제와 위험을 수반함
  - 따라서 공정한 경쟁환경을 보장하고 동일한 AI 응용 프로그램에 대한 규제 방식도 일관되도록 하기 위하여 수평적이고 교차적인 표준이 필요함
    - 한편, 특정 분야에만 해당하는 특정 위험 사항은 표준화 프로세스의 미래 AI법안의 맥락에서 적절하게 고려되어야 함. 이러한 접근방법을 통해 수평적인 표준의 개발이 다양한 산업 분야의 요구를 잘 충족시킬 수 있도록 보장할 수 있음
    - 미래 AI법을 지원하는 표준화 프로세스의 맥락에서 AI가 활용되는 특정 산업에 내재한 특정 위험을 고려하는 것은 중요함
  - (대상 범위) 기본적으로 교육 및 고용, 의료 및 헬스케어, 금융, 법 집행 및 공공 부문, 산업 자동화 및 로봇 공학 등 AI법에서 고위험군 산업으로 분류한 분야에 초점을 맞추어 분석함
    - 이외에도 소셜미디어, 콘텐츠 검열, 추천 시스템 등을 포함한 미디어 분야도 AI법과 디지털 서비스법 상 일부 의무의 대상인 점을 고려하여 해당 산업 분야를 검토함

1 교육 및 고용

- AI 기술 벤치마크에서 압도적 다수가 학습 데이터 사용에 의존하는 것으로 나타나 AI에 있어 데이터의 중요성과 윤리적 표준이 더욱 강조됨. 다양한 경제 부문에서 AI 도입이 증가함에 따라 교육과 고용에도 주목할 만한 영향을 미칠 전망임
  - 머신러닝 등 여러 인공지능의 발전으로 인해 업무의 자동화가 이루어지고, 결과적으로 고용 부문에 큰 영향을 미치고 있음[14)]

14) Tolan, S, A Pesole, F Martínez-Plumed, E Fernández-Macías, J Hernández- Orallo, and E Gómez., "Measuring the occupational impact of AI: tasks, cognitive abilit ies and AI benchmarks", Journal of Artificial Intelligence Research, 71, 191-236., (2021)

- 이러한 현상은 두 가지 메커니즘을 통해 교육 부문에 영향을 미침. 첫째, 학생들은 점점 더 자동화되는 경제와 사회를 위한 새로운 기술에 대비해야 함. 둘째, AI 및 관련 기술이 교실과 학습 시스템 수준에서 교육 과정을 변화시키고 향상시키는 데 활용될 것으로 예상됨

- 직장에서의 업무 자동화를 통해 노동 시장에 영향을 미칠 뿐만 아니라 온라인 노동 시장 및 디지털 플랫폼과 같은 새로운 형태의 업무가 급격히 증가하고, 알고리즘 관리 관행이 확대되고 있음. 이에 따라 비즈니스 성과와 근로 및 생활 환경 개선의 이점이 있음

- 한편, 알고리즘의 편향성, (성별) 고정관념의 강화, 모니터링/감시 가능성의 악용 등 인간의 일과 고용의 미래에 대한 우려도 존재함. 또한, 기계의 성능이 인간의 능력을 넘어서는 수준으로 향상됨에 따라 인간의 일자리가 위험에 처해 실업률이 크게 증가할 수 있다는 우려도 있음

교육 및 고용 부문의 한 가지 특별한 측면은 고용주와 직원, 교사와 학생 등 문화에 따라 달라지는 사회적 관계와 각 주체들이 수행하는 역할에 대해 사회적 정의가 내재해 있다는 점임

● 이러한 구조에 AI 시스템을 구현하면 투명성과 관련된 문제가 발생할 뿐만 아니라, 이러한 관계에 영향을 미쳐 권력 불균형(power imbalances)이 발생하거나 기존의 권력 불균형이 강화될 위험이 있음

- 따라서, 투명성은 AI를 위한 포괄적이고 편향되지 않으며 신뢰할 수 있는 데이터를 확보하기 위한 핵심 개념으로 인식되고 있음

● 투명성의 핵심 측면은 데이터와 알고리즘에 대한 접근성과 이러한 알고리즘의 설명 가능성이며, 데이터, 알고리즘, 교육 및 고용 알고리즘에 대해 서로 다른 투명성 기준을 설정해야 함

- 어떤 데이터가 사용되고, 데이터가 어떻게 사용되며, 이를 기반으로 어떤 의사결정을 내리는지에 대한 투명성이 보장되어야 함

- 일례로, 노동자와 학생의 데이터를 수집하여 의사결정을 내리는 과정에서 데이터 수집에 대한 통지가 원활히 이루어지지 않는 데에 대한 우려가 있음. 이는 데이터 소유권 구조 및 개인정보 침해와 관련된 문제를 낳을 수 있음

● 알고리즘의 투명성은 두 가지 문제와 관련이 있음

- 하나는 제3자의 감사를 위해 알고리즘 접근을 허용하는 것과 알고리즘의 지적 재산권 보호 사이의 균형을 찾는 것이고, 두 번째는 딥러닝 알고리즘의 복잡성이 증가함에 따라 비전문가가 알고리즘을 해석하는 데 어려움을 겪는 문제를 다룸

- 알고리즘에 의해 내려진 결정에 이의제기를 어렵게 하는 불투명성이 일반화되어 있음. 알고리즘 규칙에 대한 투명성을 확보하려면 직장과 학교에서 AI 시스템의 명확한 역할 경계가 필요함

- 전문가가 아닌 일반인은 알고리즘을 제대로 이해하지 못하여 권력의 불균형이 발생할 수 있으므로,

데이터의 수집 및 저장, 알고리즘 사용 과정에서 이해관계자가 항상 지켜볼 수 있도록 해야 함

이러한 AI 시스템의 불투명성과 복잡성 문제를 해결하기 위한 몇 가지 방안은 다음과 같음

- 모든 이해관계자가 사례별로 알고리즘 결정에 대한 설명을 이해할 수 있도록 개인의 권리를 보장해야 함
  - 이 설명에는 어떤 데이터를 검토했는지, 특정 측면에 가중치를 부여했는지, 특정 결론에 어떻게 도달했는지에 대한 정보가 포함되어야 함
- 잠재적으로 소외된 집단의 문제를 해결하기 위해서는 AI 도구가 소외 집단을 어떻게 대하는지에 대한 이해가 필요함. 이는 데이터세트의 대표성과 연관이 있으며, 합리적이라면 조정될 수 있음
- 권력의 불균형 문제를 해결하기 위해 몇 가지 데이터 및 AI 거버넌스 메커니즘이 제안되었는데, 여기에는 데이터 상호 우호성, 투명성 및 설명 가능성에 대한 권리가 포함되어 있음. 즉, 직원, 근로자, 학생은 자신에게 영향을 미치는 AI에 사용되는 데이터에 접근할 수 있는 권리를 가져야 함
  - 또한, 고용주와 직원, AI 제공업체, 학교, 교사 간의 사회적 대화 메커니즘이 제안되었으며, 직원과 학생은 대표를 통해 직장과 학교에서 알고리즘 구현을 공동 결정할 권리를 부여하고 편향성 평가 및 설명 가능성에 대한 외부 전문가 감사를 지원받아야 함
  - 통제와 규제를 달성하기 위한 또 다른 방안인 오픈 소스 모델은 투명성과 열린 대화를 기반으로 하는 분산형 알고리즘 평가방식으로, 노동자와 학생이 AI 시스템 기반의 의사결정 과정에 데이터 보호 규정에 따른 통제된 방식으로 참여할 수 있으며, 외부 데이터 트러스트[15]를 마련하는 방안도 제시됨

AI 시스템용 데이터의 개발 단계를 고려한 과제와 해결책을 표준화 개념에 매핑하는 작업을 수행하면서 가장 중요하고 시급한 문제가 용어 정립과 측정임

- 편견과 차별(기술 맥락에서 서로 다른 의미로 사용됨), 투명성, 정확성 등 AI 및 데이터 관련 문제에 대한 명확하고 공통적인 용어가 제대로 정립되지 않아 분야 간, 부문 간 협업이 제대로 이루어지지 않음
  - 예를 들어, 투명성이라는 용어도 데이터의 포괄성, 비편향성, 신뢰성 제고를 위해 필수적인 핵심 개념으로 인정되지만 용어 자체가 다의적인 개념이기 때문에 정확한 정의가 필요함
  - 따라서 매우 중요한 과제는 편향 대 차별(Bias vs. Discrimination), 투명성 및 정확성과 같은 데이터 문제에 대해 명확한 공통 용어를 설정하는 것임
  - 노동자 또는 학생의 성과를 평가할 때 AI 시스템 단독으로 평가하지 않도록 기준 설정이 중요함
  - AI법은 AI 시스템 규제를 위험 기반 접근 방식으로 모범사례를 제시하고 있으므로 이를 채택할 필요가 있으며, 동시에 딥러닝 알고리즘의 블랙박스 특성을 이해하기 위해 개인화된 설명이

15) 외부 데이터 트러스트(external data trusts): 제3자가 개인들의 데이터 또는 데이터 권한을 관리하는 것

필요함

| 표 4 | 교육 및 고용을 위한 선별된 표준화 범주와 AI 가치 사슬에 따른 사전 규범 연구 및 표준 요구사항 개요

	용어	측정	성능 특성	호환성	규제 평가
AI 시스템 배포과 마케팅 - 규제 평가 - 사용자 - 투명성 및 명세서 - 책임 및 의무 - 유지 보수, 시장 후속 조치 및 편향성 모니터링 - 공급망			- 자율성 결여 - 개인정보보호와 개방성의 균형 - 인공지능 신뢰도	- 의사결정 투명성 - 일반 사용자 지침 - 사용자 간 표준화 - 사용자 교육	- 설명 가능성에 달린 배포 조건 - 기술 유지 관리
AI 시스템 생성 및 생산 - 데이터 세트 및 알고리즘 (편향성 포함)/모델 - 사이버보안 - 시스템 설계 및 통합 - 업스케일링 및 평가 - 품질 관리	- 개인정보보호 - 투명성	- 개인화된 설명 - 복잡성	- 숨겨진 차별 - 오픈소스 모델 - 정보 비대칭에서의 권력 불균형	- 도메인 전문성 구현 - 데이터 사용에 대한 투명성	- 참여형 평가
데이터 생성 - 컴파일링, 준비, 편향성 검사 - 분석, 가공, 라벨링 - 라이선스 및 제한사항 - 공유 및 마케팅	- 명확한 공통 용어 없음 - 근로자에 관한 데이터 명확도 - 모호하고 유동적인 용어 - 교육 관측 가능성[16](모델) vs 제약사항 및 중요성 - 이해관계자 식별	- 데이터 내 소수집단 대표성 - 정확도 - 설정된 메트릭 없음 - 다각적 관점/소스 데이터 협업 - 주요 매개변수 식별 위한 시사점 연구	- 교육 관측 가능성(모델) vs 한계 및 중요성 - 데이터 표현의 정확성 및 완전성		- 외부 데이터 신뢰

2 법 집행 및 공공 부문

법 집행 및 공공 부문에서 AI를 사용하는 것은 범죄자 체포, 법원 판결, 시민 감시, 공공 보조금 회수 등 법적·사회적으로 큰 결과를 초래할 수 있기 때문에 특히 민감한 부문임

- 법 집행 및 공공 부문에서는 특히 데이터에 초점을 두고 신뢰할 수 있는 AI 시스템을 개발하기 위해 어떠한 표준을 도입하고 해결책을 수행할지 분석함

16) AI 관측 가능성은 기계 학습 시스템의 모든 부분에서 통계, 성능 데이터 및 지표를 수집하는 것을 뜻함

● 유럽위원회의 AI 규제안에 따르면 원격 생체 인식, 범죄 발생 예측, 감정 상태 감지, 범죄 분석과 같은 분야에서 AI의 사용은 높은 위험을 수반함[17)]

- 올바른 법 집행을 위해 AI 시스템을 학습하고 평가하기 위해서는 양질의 데이터를 확보하는 것이 필수적이며, AI 시스템 수명주기 동안 적절한 거버넌스 관행을 보장하는 것도 중요함

법 집행 및 공공 부문에서 시급하게 해결해야 할 과제는 다음과 같음

- **실제 데이터와 실험실 데이터 격차:** 법 집행용 AI 도구 배포에는 실시간 데이터 캡처(카메라, 스캐너, 지문 판독기 등), 환경(조명, 기상 조건 등), 사람(민족성, 출신지, 연령, 성별, 역량 등의 관점) 등 매우 다양한 센서가 사용되므로 모든 시나리오를 고려하기 어려움. 또한, 개인정보 보호 문제 및 데이터 부족과 같은 문제로 인해 대량의 실제 데이터를 수집하는 것이 불가할 때가 많음. 이에 시뮬레이션된 데이터, 실험실에서 임시로 생성된 데이터를 기반으로 시스템 개발을 하는 것이 일반적이나, 실험실에서 운영 환경으로 이전하는 과정에서 정확성, 신뢰성이 보장되지 않는 경우가 있음

- **인구통계학적 편향 문제:** 이는 법 집행 및 공공 부문에서 특히 시급한 문제로, 데이터 편향이 주로 여성, 노인, 어린이에게 나타나고 모든 인종의 데이터가 반영되지 않기 때문에 데이터 불균형이 나타남. 또한, 데이터 수준에서 인구통계의 모델링 방식에 대한 합의가 없어 편향성을 측정하는 작업도 어려움

- **신뢰할 수 있는 출처에서 데이터 수집:** 데이터 수집 과정은 시간이 많이 소요되기 때문에 AI 개발자는 제3자가 제공하는 데이터세트에 의존하는 경우가 많음. 데이터 세트를 공급하는 민간 기업에서 데이터세트를 구입하는 것이 점차 일반화되고 있는데, 이는 현재 데이터의 출처와 신뢰성을 인증하는 규정이나 표준 수단이 없어 우려스러운 현상임

- **설명 가능성 부족:** '블랙박스' 접근 방식으로 AI의 사용이 증가함에 따라 견고성과 정확성이 향상되는 이점이 있지만, 설명 가능성이 부족해 여기에 대한 더 많은 연구가 필요함

- **법 집행 맥락에서의 AI 이해:** 법 집행 주체인 경찰과 법원은 아직 AI 활용의 장단점에 대해서 잘 인지하지 못하고 있으며, 법 집행 주체에 대한 AI 관련 교육이 제공되어야 함

법 집행 및 공공 부문에서 신뢰할 수 있는 AI 채택을 위한 여러 표준화 기회가 있음

● 현재 생체 인식 애플리케이션은 개인의 식별(얼굴, 지문, 홍채 인식)에만 초점을 맞추고 있기 때문에 생체 인식을 넘어 다른 인공지능 기반 기술(재범 예측, 금융 사기, 허위 정보 통제, 자동 음성 전사 등) 도입이 필요함

- 법 집행 기관을 위한 데이터 공간을 마련하면 서로 다른 법 집행 기관뿐만 아니라 국내 및 국가 간 데이터 공유를 원활히 수행할 수 있음

- 해당 분야에서 특히 중요한 문제는 편향성으로, AI 시스템의 수명주기 동안 편향성을

17) European Commission, 「European Commission's proposal for a Regulation on Artificial Intelligence」, Luxembourg: European Commission, (2021)

모니터링하기 위한 한계선과 지표(redlines and metrics)에 대한 합의에 도달해야 편향성 문제를 완화할 수 있음

- 해당 부문에 대한 표준에서 다루어야 할 두 가지 고려사항은 1) 표준은 주로 중소기업이 채택할 것이라는 점과 2) 표준에는 시스템의 주요 사용자(LEA(law enforcement agency, 법 집행기관), 판사, 배심원)를 위한 AI 및 인권 교육에 대한 프로세스가 포함되어야 하는 점임

| 표 5 | 법 집행 및 공공 부문을 위한 선별된 표준화 범주와 AI 가치 사슬에 따른사전 규범 연구 및 표준 요구사항 개요

	용어	측정	성능 특성화	호환성	규제 평가
AI 시스템 배포과 마케팅 - 규제 평가 - 사용자 - 투명성 및 명세서 - 책임 및 의무 - 유지 보수, 시장 후속 조치 및 편향성 모니터링 - 공급망			- 운영 테스트 방법 및 벤치마크		- 법 집행 기관에 인권 문제 교육 제공 - 경찰과 판사에게 인공지능 교육 제공 - 중소기업 고려 (비례성)
AI 시스템 생성 및 생산 - 데이터 세트 및 알고리즘 (편향성 포함)/모델 - 사이버보안 - 시스템 설계 및 통합 - 업스케일링 및 평가 - 품질 관리	- 편향과 관련한 용어에 대한 공통 분류	- 데이터 및 모델의 편향성 평가를 위한 지표 및 벤치마크		- 상호운용성에 대한 표준 - 통신 및 사이버 보안에 관한 표준	- 증거 및 설명 가능성에 대한 기준
데이터 생성 - 컴파일링, 준비, 편향성 검사 - 분석, 가공, 라벨링 - 라이선스 및 제한사항 - 공유 및 마케팅		- 데이터 편향성에 대한 표준화된 메트릭	- 신뢰할 수 있는 조직에서 데이터 확보	- 법 집행 기관을 위한 표준화된 데이터 공간	- 데이터 공급업체 규제

3 금융

- 금융용 AI의 잠재적인 문제로는 알고리즘 편향으로 인해 차별적인 의사결정을 내려 소비자에게 여러 악영향을 미칠 수 있음
  - 금융 서비스는 소비자를 보호하고, 효과적인 경쟁을 촉진하며 금융 시스템의 무결성과 복원력을 강화해야 함
    - 금융용 AI에는 다양한 수준의 위험이 존재함. 필수적인 민간 및 공공 서비스(예: 신용 평가가 시민의 대출 기회를 거부하는 것)와 같은 높은 위험부터 투명성 의무가 있는 특정 AI 시스템(예: 금융 정보에 대한 AI 시스템 사용 시 사용자가 기계와 상호 작용하고 있다는 사실을 인지하여 정보에 입각한 결정을 내릴 수 있어야 하는 제한적 위험)에 이르기까지 다양함
    - 안전한 AI 도입을 위해 잠재적인 이점과 해악에 관한 논의, 현행 규제 프레임워크의 적용 방식과 지원 방안, 규제 당국 간 지식 공유와 협업, 이해관계자의 의견을 공유할 수 있는 기회가 필요함
  - 데이터 품질 모델 및 측정에 대한 표준은 이미 존재하지만, AI 적합성은 직접적으로 평가되지 않음. 그러므로 데이터 품질 모델은 AI에만 국한되어서는 안 되고, AI에 사용되는 데이터의 사용 전반에 대한 논의를 필요로 함. 따라서 품질 모델은 AI와 관계없이 편향성이 의사 결정에 어떻게 영향을 미치는지 다양한 단계에서 모델 결과를 테스트하여 올바른 결과를 얻어야 함
- 금융 산업의 데이터 품질 문제를 다루기 위해서는 먼저 데이터 품질 및 모델 품질에 대한 책임 메커니즘을 다루는 공통적인 접근 방식의 필요성과 표준 설정 목적(공정성 및 비편향성을 확인하기 위한 표준, 지침 모범 사례)을 위한 용어를 정의하고 알고리즘 감사를 고려할 필요가 있음

| 표 6 | 금융 부문을 위한 선별된 표준화 범주와 AI 가치사슬에 따른 사전 규범 연구 및 표준 요구사항 개요

	용어	측정	성능 특성화	호환성	규제 평가
AI 시스템 배포과 마케팅 - 규제 평가 - 사용자 - 투명성 및 명세서 - 책임 및 의무 - 유지 보수, 시장 후속 조치 및 편향성 모니터링 - 공급망					

	용어	측정	성능 특성화	호환성	규제 평가
AI 시스템 생성 및 생산 - 데이터 세트 및 알고리즘 (편향성 포함)/모델 - 사이버보안 - 시스템 설계 및 통합 - 업스케일링 및 평가 - 품질 관리	- 빅데이터는 데이터 품질 저하의 위험을 증가시키며, 이러한 위험은 건전성 리스크로 이어질 수 있음 - 공정성 정의와 맥락상 공정성이 의미하는 바에 대한 정의	- 모든 부문에 적용 가능한 데이터 품질 모델 - 기존 표준에 기반한 공통 접근 방식	- 모델 결과 검증 vs. 데이터 품질 점검 - 활용 대상에 따른 맞춤형 모델 필요		
데이터 생성 - 컴파일링, 준비, 편향성 검사 - 분석, 가공, 라벨링 - 라이선스 및 제한사항 - 공유 및 마케팅	- AI/금융 도전과제 식별	- 데이터 대표성 - 편향의 원인이 되는 학습 데이터	- 데이터 출처 및 특성		- 데이터/모델 품질에 대한 책임

4 소셜 미디어, 콘텐츠 조정, 추천 시스템을 포함한 미디어

- 소셜 미디어, 콘텐츠 조정(moderation), 추천 시스템을 포함한 미디어용 AI는 AI법의 이니셔티브와 유럽위원회가 제안한 이니셔티브인 디지털 서비스법(DSA) 및 디지털 시장법(DMA)이 연결되어 있음
  - DSA와 DMA에는 아래 두 가지 목표가 있음
    1) 디지털 서비스를 이용하는 모든 사용자의 기본권이 보호되는 보다 안전한 디지털 공간을 조성하고자 함
    2) 유럽 단일 시장과 전 세계에서 혁신, 성장 및 경쟁력을 촉진하기 위한 공정한 경쟁의 장을 구축하고자 함

- 미디어, 미디어에 사용되는 소셜 미디어 데이터 또는 콘텐츠 조정 및 추천을 위해 소셜 미디어를 사용하는 AI 애플리케이션의 경우, 데이터 품질과 모델 학습에 사용되는 데이터의 편향성을 평가하는 것이 중요함
  - 콘텐츠 조정(권장)은 유해한 콘텐츠, 가짜 콘텐츠 및 기타 부정적인 유형의 콘텐츠를 제거하는 데 도움이 됨. 추천 시스템은 대량의 정보를 선별하여 사용자와 관련된 콘텐츠를 찾아내는 데 도움이 되지만, 단일 관점에만 집중하게 되어 사용자가 잠재적으로 이용할 수 있는 정보가 줄어들 수 있음

- 현재 소셜 미디어 플랫폼의 비즈니스 모델은 사용자의 관심을 장기간 확보하고 유지하는 것이 관건이며, 이러한 맥락에서 제안되는 콘텐츠는 더욱 감정적이고 양극화되어 급진화, 허위 정보의 확산과 같은 매우 위험한 결과를 초래할 수 있음

● 소셜 미디어 데이터와 대량의 웹 데이터는 자연어처리(NLP) 딥러닝 모델을 학습하는 데 사용됨. 이러한 데이터는 콘텐츠와 콘텐츠에 기여하는 사용자의 이질성 및 고유한 주관성으로 인해 성별, 정치적 성향, 인종, 종교 등 다양한 유형의 편향성이 나타날 수 있음

AI 시스템이 미디어 콘텐츠에 적용될 때, AI의 편향성 개념을 정의할 필요성과 콘텐츠 조정 및 추천 시스템에 사용되는 데이터의 특성을 이해하는 것이 중요함

● 후자의 경우, 사용자의 개인적 특성이 시스템의 품질에 가장 중요하고 AI 시스템이 개인적 선호도에 편향되는 것이 바람직하지만, 전자의 경우 편향성은 양극화와 정보 거품 효과로 이어질 수 있음

● 소셜 미디어 플랫폼을 소유한 기업 간에는 다른 업계 플레이어와 학계 및 연구 기관에 비해 데이터에 대한 접근 권한에 큰 차이가 있다는 것이 밝혀짐

- 불평등한 접근은 소규모 플레이어가 강력한 AI 모델을 만들지 못하게 할 수 있음

- 벤치마킹 목적으로 사용할 수 있는 관련 데이터세트에 대한 접근성이 부족하면, 모델을 학습하고 테스트할 수 있는 관련 데이터가 없기 때문에 편향성 여부를 평가하는 작업이 매우 어려워질 수 있음

개인과 집단의 공정성 사이의 연관성과 편향성을 정량화하기 어렵다는 점에 특히 주의를 기울여야 하며, 도메인 과제에 의존하지 않고 일반적인 방식으로 문제를 해결하는 것이 중요함

● 데이터에 존재하는 잠재적 편향성을 이해하고, 이를 체계적으로 정의하며, 감지하고 완화할 수 있는 적절한 방법을 개발하기 위해서는 관련 데이터세트에 대한 접근이 핵심임

- 이 프로세스는 일반데이터보호규정 및 개인정보보호 규정에 따라 수행되어야 하며, 개념이 흐트러지지 않도록 적절한 정의가 필요함

● 전체 연구 커뮤니티가 사용할 수 있고 다양한 유형의 편향성 테스트를 위해 벤치마킹 데이터세트의 생성이 하나의 해결책이 될 수 있음

● 표준화 기회 중 우선순위가 높은 4가지의 중요 과제는 다음과 같음

① 미디어용 AI 시스템에 사용되는 데이터 품질과 공정성에 관한 용어 정의
② 데이터세트 품질을 평가할 체크리스트 정의
③ 미디어 환경에서 사용되는 AI 시스템의 학습, 테스트 및 감사를 위해 양질의 데이터에 대한 개방형 데이터 액세스를 보장하는 메커니즘 개발
④ 미디어 환경에서 사용되는 AI 애플리케이션의 편향성 감지를 위한 개방형 벤치마킹 데이터세트 및 모델 생성

| 표 7 | 미디어 산업을 위한 선별된 표준화 범주와 AI 가치사슬에 따른 사전 규범 연구 및 표준 요구사항 개요

	용어	측정	성능 특성화	호환성	규제 평가
AI 시스템 배포과 마케팅 - 규제 평가 - 사용자 - 투명성 및 명세서 - 책임 및 의무 - 유지 보수, 시장 후속 조치 및 편향성 모니터링 - 공급망	- 미디어 애플리케이션에 대한 AI 및 데이터의 적합성 정의				- 데이터 기반 비즈니스를 위한 새로운 비즈니스 모델 정의
AI 시스템 생성 및 생산 - 데이터 세트 및 알고리즘 (편향성 포함)/모델 - 사이버보안 - 시스템 설계 및 통합 - 업스케일링 및 평가 - 품질 관리	- 주요 애플리케이션에서 집단 대 개인의 공정성 정의 - 애플리케이션 사용의 공정성 정의	- 소셜 미디어 플랫폼 및 기타 관련 데이터 소스의 데이터 접근 보장	- 편향 탐지 및 완화를 위한 벤치마킹 모델을 생성		- 감사를 위한 벤치마킹 모델을 연구단체에 공개
데이터 생성 - 컴파일링, 준비, 편향성 검사 - 분석, 가공, 라벨링 - 라이선스 및 제한사항 - 공유 및 마케팅	- 미디어용 AI에서 데이터 품질에 대한 용어 정의 - 미디어용 AI의 데이터 품질 측면 정의	- 전체 연구단체를 대상으로 교육 및 조정이 가능한 벤치마킹 데이터세트 생성 - 데이터 출처 및 데이터 품질 (규정 준수, 개인정보 보호 측면 대표성)에 부합하는 라벨링 시스템 생성			- 표준화 및 감사 목적으로 사용할 관련 데이터에 대한 접근 메커니즘 마련

5 의료 및 헬스케어

- 표준화 활동은 의료 및 헬스케어 분야에서 AI의 잠재적 기회를 포착하는 데 중요한 역할을 하며, 신뢰성 문제를 해결하고 환자와 사용자의 안전을 보장함
  - 표준화 기관, 과학계 및 규제 커뮤니티 등 다양한 커뮤니티를 연결하는 것은 의료와 같은 다학제적 분야에서 유용하고 쉽게 적용할 수 있는 표준을 개발하는 데 핵심적인 요소임

- 의료 분야에서 사용되는 AI 시스템의 위험 평가를 위한 명확한 프레임워크는 부문별 규제에 따른 요건을 명확히 하고 계층화하여 혁신적인 접근 방식을 지원하는 데 도움이 됨
- AI 시스템의 기반이 되는 통계 및 데이터 처리가 애플리케이션과 독립적이라는 점을 고려할 때, 표준화는 우선 데이터 품질 요건, 편향성 방지, 명료성 등 신뢰성과 같은 수평적 표준에 초점을 맞춰야 함
- 이와 대조적으로 수직적 지침은 특정 산업 및 사용 사례 또는 애플리케이션의 요구사항을 다룰 수 있음
- 또한, 표준, 보고서 및 지침 문서에 대한 실용적인 로드맵 수립이 중요한데, 이는 표준화 커뮤니티가 규제 커뮤니티와 긴밀한 대화를 통해 개발해야 함

의료 분야에서 AI 시스템을 배포할 때 편향성을 주요 과제 중 하나로 지적하는 과학 문헌이 점점 더 많아지고 있음. 편향성 문제는 AI 지원 의료 솔루션의 수용성을 저해할 수 있음

- 이 편향성 문제에는 암 치료나 코로나 19 진단을 위한 AI 시스템에서 의료 불평등을 초래하는 편견이 해당함. 따라서 편견을 피하기 위한 명확한 프레임워크를 제공하는 표준은 AI의 안전한 사용을 위한 핵심 원동력이 될 것임
- 편향성 문제 해결과 관련하여 현재 진행 중인 내용에는 TRIPOD-AI 리포팅 가이드라인과 AI를 사용한 진단 및 예후 예측 모델에 대한 편향성 위험 도구인 PROBAST-AI가 있음

- TRIPOD-AI는 편향을 설명하고 줄이는 데 일관된 접근 방식을 제공하기 위해 제안된 체크리스트를 포함하며, 잔존 편향을 투명하게 전달할 수 있음
- AI에 사용되는 데이터와 알고리즘에 대한 '품질 라벨'을 검토하고 채택하는 것도 실용적인 해결책이 될 수 있음

편향성, 대표성 및 견고성을 평가하기 위한 데이터는 대부분 식별이 가능하며 국제적으로 다양한 데이터 보호 표준을 적용받음

- 그러나 개발자 및 제공업체는 환자나 의료기관에 있는 데이터에 접근하기 어려워 데이터 접근 권한을 제공할 수 없는 경우가 많음에도 불구하고, EU의 AI법 초안은 관련 당국이 전체 데이터세트에 접근할 수 있도록 요구하고 있음

- 제3국 데이터 보호법이 데이터 전송을 금지할 수 있음. 따라서 AI법의 요건은 EU에서 생산된 AI 시스템이 해당 데이터를 사용할 수 없게 되고 개발자는 견고성에 부정적인 영향을 미칠 수 있는 대표성이 낮은 데이터세트(하위 집단, 희귀질환 등)에 의존할 수 밖에 없는 결과를 초래할 수 있음

- 데이터에 알고리즘을 적용하는 것도 기술적 한계를 고려할 때 이상적이지 않을 수도 있음

- 표준화 우선순위와 관련하여 데이터 사용과 데이터세트에 대한 액세스 제공의 필요성에 대해 균형 잡힌 접근이 필요하며, 특히 민감한 데이터의 경우 데이트 측정 기준을 설정하고

데이터를 설명하는 방법에 대한 표준이 있어야 함

- 제공자가 데이터에 직접 액세스할 수 없는 경우, 관련 제3자는 표준에 따라 데이터에 대한 충분히 상세한 설명을 제공할 수 있어야 하며, 편향성 및 차별을 식별하고 이에 대한 교육을 제공하며 사용자에게 발생한 편향성에 대해 경고할 수 있어야 함

● 제조업체는 의도한 목적(질병 및 환자 집단, 사용된 데이터세트의 크기 등)에 대한 데이터의 적합성을 입증하는 통계적 방법을 사용하여 데이터 프로파일링을 수행해야 함

- 이러한 '품질 카탈로그'를 해결하기 위해 측정 방법을 데이터 관리 문서에 요약하여 AI 모델 개발 프로세스의 핵심 요소를 파악하고, 이 문서가 품질 관리 시스템의 일부가 되어야 함

● 새로운 AI 표준은 모순되는 요구사항과 지침으로 인한 중복 또는 혼란을 피하기 위해 보건 분야에서 이미 확립된 표준의 환경에 맞아야 함

- 의료기기 분야에서는 예를 들어 소프트웨어의 안전 분류와 관련하여 IEC 62304 및 IEC 82304-1과 같은 관련 표준을 고려해야 함. 혁신적인 접근 방식을 유지하기 위해 MDR 또는 IVDR에 해당하는 AI 시스템에 대한 AI 표준은 각 법률에 명시된 요구사항을 강화하지 않으면서도 명확해야 함

본 워크숍에서는 표준화를 위해 기술적 측면의 4가지의 요구사항을 확인함

- **용어**: 지역사회와 전문가, 표준 사용자 간의 효과적인 의사소통을 위해 포괄적인 용어가 정립되어야 하며, 용어는 데이터 생성부터 AI 시스템 구축 및 마케팅에 이르는 개발 경로의 모든 요소를 다 다룰 수 있어야 함

- **데이터 품질 측정 지침**: 데이터의 품질과 적절성을 측정하기 위해 적절한 지침 및 표준을 개발해야 하며, 가능한 경우 디바이스의 개인화가 가능한 측면이 포함되어야 함. 이는 개인의 필요와 매개 변수에 맞게 기기를 맞춤화해야 하는 보건 분야와 특히 관련 있음

- **AI 시스템의 필수 요소를 문서화하는 내재 방법 지침**: 시스템 내 적용 가능한 영역과 한계, 잠재적 위험 등을 포함한 안내서는 AI 시스템 개발자가 인공지능 시스템의 필수 요소를 확인할 수 있도록 함

- **안전 및 사이버보안에 대한 지침**: 의료목적으로 사용되는 AI 시스템의 안전성, 유효성, 위험성을 평가하는 방법에 대한 지침이 필요하며, 사이버 공격에 대한 취약성 평가도 포함되어야 함

AI를 활용하면 반복적인 실무를 자동화함으로, 의료 업계 종사자는 환자와 더욱 상호 작용을 늘리고 환자 진단 및 치료에 더 집중할 수 있음. 또한, AI 시스템은 환자의 건강을 개선하고 삶의 질을 크게 향상시킬 수 있는 잠재력을 가진 디바이스의 개인화를 가능하게 할 수 있음

● 본 워크숍에서는 1) 의료 커뮤니티 내에서 전반적으로 활용할 수 있는 용어 및 개발자와 제조업체의 표준 채택 필요성, 2) 데이터 품질을 보장하기 위해 적합성에 대한 지표 및 통계적 측정 등 AI 모델의 핵심 요소를 설명하는 가이드라인, 3) 의료 관련 표준에 대한 필요성을 파악하는 관점에서 수평적 표준 분석을 최우선 과제로 확인함

| 표 8 | 의료 및 헬스케어를 위한 선별된 표준화 범주와 AI 가치사슬에 따른 사전 규범 연구 및 표준 요구사항 개요

	용어	측정	성능 특성화	호환성	규제 평가
AI 시스템 배포과 마케팅 - 규제 평가 - 사용자 - 투명성 및 명세서 - 책임 및 의무 - 유지 보수, 시장 후속 조치 및 편향성 모니터링 - 공급망	✓	- 최신 데이터 측정 프로토콜, 데이터 지표 및 관련 임계값 설정 방법		- 배포/판매된 AI 시스템 문서화 지침(적용가능성, 제한사항, 위험성, 데이터 설명, 데이터 위생 인증서[18])	- 배포/판매된 AI 시스템 문서화 지침(적용가능성, 제한사항, 위험성, 데이터 설명, 데이터 위생 인증서)
AI 시스템 생성 및 생산 - 데이터 세트 및 알고리즘 (편향성 포함)/모델 - 사이버보안 - 시스템 설계 및 통합 - 업스케일링 및 평가 - 품질 관리	✓	- 최신 데이터 측정 프로토콜, 데이터 메트릭 및 관련 임계값 설정 방법	- 안전 및 사이버 보안 측면		
데이터 생성 - 컴파일링, 준비, 편향성 검사 - 분석, 가공, 라벨링 - 라이선스 및 제한사항 - 공유 및 마케팅	✓	- 최신 데이터 측정 프로토콜, 데이터 메트릭 및 관련 임계값 설정 방법	- 안전 및 사이버 보안 측면		

6 산업 자동화 및 로봇 공학

- 로봇을 구동하는 머신러닝 알고리즘은 빠르게 발전하고 있음. AI 소프트웨어 및 시스템 제공업체는 태그가 지정된 예제 수를 줄이는 등 알고리즘이 학습할 수 있도록 하기위해 반지도 또는 자가지도 방식으로 개발 노력을 기울여왔음
  - 산업 자동화 및 로봇 공학 분야는 표준화와 관련하여 특별한 주목을 받고 있으며, 로봇 공학에 특화된 여러 표준화 이니셔티브가 존재함

18) 데이터 위생(Data Hygiene)은 일반적으로 데이터가 깨끗하다는 것을 보장하는 모든 지속적인 프로세스를 의미함. "깨끗한"이란 주로 오류가 거의 없는 데이터를 의미하며, 역으로 더러운 데이터는 오류를 포함하는 데이터로, 오래된 데이터, 불완전한 데이터, 중복된 데이터 또는 단순히 정확하지 않은 데이터를 포함. 이러한 오류는 데이터가 시스템 내에서 어떤 시점에서든 도입될 수 있으며, 초기에 잘못 입력되었을 수도 있고 레코드를 업데이트할 때 실수로 변경되었을 수도 있음. 깨끗한 데이터의 필요성을 고려할 때, 지속적인 데이터 위생 관행을 유지하는 것이 중요함

- 예를 들어, ISO/TC 299 로보틱스 표준 시리즈에는 현재 37개의 표준이 발표 또는 개발 중임[19]
- 그러나 포괄적이고 편향되지 않으며 신뢰할 수 있는 로봇 시스템을 위해서는 시스템 학습용 데이터와 물리적 환경에서 작동하는 로봇의 수집 데이터의 품질을 관리해야 함

인간과 로봇의 상호작용의 관점에서, 산업 자동화 및 로봇 공학의 표준화와 관련하여 세 가지 주요 과제가 있음

- 첫째, 데이터 수집에 대한 특별한 주의가 필요함
  - 예를 들어, 개인용 로봇은 로봇과 직접 상호 작용하는 사람에 대한 데이터뿐만 아니라 주변 사회 환경의 데이터도 수집할 수 있음. 이러한 경우 시스템의 필요에 따라 데이터를 선별할 수 있는 메커니즘이 필요하며, 삭제할 수 있는 시스템도 설계해야 함
- 둘째, 데이터 공유에 대한 관리가 중요함
  - 개인용 로봇이 특정 인간과 상호작용하는 동안 수집하는 데이터는 민감한 정보일 수 있으며, 인간은 이를 공유하지 않으려 할 수도 있음. 따라서 데이터 수집 시 맥락을 고려해야 함
  - 로봇이 인간을 조작할 수도 있는데, 이는 어린이와 같은 취약계층에 특히 우려되는 부분임. 이를 해결할 수 있는 전문가 풀을 구성하기 위해 전문 학계의 교육이 필요하며, 전문 교육은 투명성을 극대화하기 위한 자격증 취득으로도 이어질 수 있음
- 셋째, AI 기반 소셜 로봇과의 상호작용에 대한 일반 대중의 교육이 필요함
  - 이는 공교육을 통해서도 이루어질 수 있으며, 일반 대중이 최소한의 AI 및 로봇 공학 지식을 갖출 수 있도록 보장할 수 있음

산업 자동화 및 로봇 공학 표준화에 대한 인식이 부족하다는 우려가 제기되고 있으며, 이에 따라 로봇 공학 표준을 위해 모든 AI에 쉽게 접근할 수 있는 포털의 중요성이 제기되고 있음

- 로봇 공학에서 안전은 데이터 기반 방식에 대한 비교 가능한 성능, 안전 메트릭과 안전 인증서(동작 증명)가 필요하기 때문에 테스트 및 인증이 어려운 과제임
  - 따라서, 로봇 공학을 위한 모델 없는 데이터 기반 AI 방법에는 시뮬레이션 테스트와 실제 접근 방식과의 비교가 필요하다는 의견이 제안됨. 이를 위해 표준 시뮬레이터와 테스트 환경을 갖춘 인프라가 필요하며, 이를 통해 자동화된 테스트 접근 방식을 용이하게 할 수 있음
  - 또한, AI 기반 센서를 사용하는 로봇 공학의 안전과 관련하여 기존 표준의 경계를 넘어서는 모델 기반 안전 검증으로, '마이크로' 동작을 실행하기 전에 즉각적으로 인증하는 '스마트 안전'의 도입이 제안됨. 이를 위해서는 AI 기반 안전 센서의 신뢰성을 검증할 수 있는 명확하고 강력한 지표가 필요하며, 위험 평가와 같은 모델 기반의 접근 방식이 더욱 발전해야 함

19) https://www.iso.org/committee/5915511/x/catalogue/

AI 시스템의 가치사슬은 데이터 생성, 모델 개발, 시스템 배포 및 운영 총 3단계로 분류됨

- AI 시스템의 가치사슬 분류 시 첫 번째 단계가 데이터 생성임
  - 인간 환경을 탐색하는 물리적 로봇은 처음 의도한 목적과 관련 없는 데이터를 수집하여 프라이버시 문제를 제기할 수 있으며, 이를 위해 데이터 정제 및 큐레이션 기술뿐만 아니라 유럽 일반 개인정보보호법(GDPR)과 같은 기존 규정에 따라 불필요한 데이터를 제거하는 방법이 필요함
  - 추가 표준 및 기술 사양은 AI 제공업체가 데이터 편향성에 대해 문제를 해결할 수 있도록 효과적인 완화 조치를 적용하고, 일반적으로 제외되는 인구 집단 데이터의 대표성을 보장할 수 있도록 지원해야 함
- AI 시스템의 가치사슬 두 번째 단계인 모델 개발 과정에서 로봇 시스템의 견고성 및 신뢰성 검증을 위한 메트릭이 필요하다는 점이 지적되었음
  - 이를 위해 AI 로봇 공학을 위한 검증 프레임워크를 개발하고 활용해야 하며, 프레임워크를 개발하기 위해 공통 용어가 필요함
- 가치사슬 세 번째 단계로 실제 환경에 AI 시스템을 배포하는 것과 관련하여, AI 시스템은 전체 수명주기에 걸쳐 위험 평가, 감독을 포함한 모든 기술적 요구사항을 고려해야 함
  - 이 단계에서 실제 사용자는 로봇을 과도하게 의존하거나 반대로 신뢰하지 않을 수 있기 때문에 시스템의 투명성과 설명 가능성을 보장하는 표준을 도입한다면 이러한 문제를 해결할 수 있음
  - 예를 들면, 1) 개념, 데이터 생성부터 모델 학습, 시스템 개발, 배포 및 운영까지 AI 수명 주기 동안 적용해야 하는 프로세스로서의 위험 관리 및 평가 2) 정확성, 견고성, 보완 및 안전에 대한 검증 방법도 개발 프로세스의 각 단계에 맞게 조정되어야 함 3) 전반적인 표준화 환경과 관련하여 적절한 표준을 탐색하고 공동으로 채택하는 데 도움이 되는 지침이 필요함

| 표 9 | 산업 자동화 및 로봇 공학을 위한 선별된 표준화 범주와 AI 가치 사슬에 따른 사전 규범 연구 및 표준 요구사항 개요

	용어	측정	성능 특성화	호환성	규제 평가
AI 시스템 배포과 마케팅 - 규제 평가 - 사용자 - 투명성 및 명세서 - 책임 및 의무 - 유지 보수, 시장 후속 조치 및 편향성 모니터링 - 공급망	- 안전성을 평가하기 위한 기술 지식 부족 - 안전요구사항에 대한 체크리스트 - 인력 교육	- 표준 지침의 식별	- 투명성 및 문서화 - 취약 계층 - HRI[20]의 영향	- 위험 관리 및 평가 - 새로운 프레임워크 및 하이브리드 표준화 - 다른 표준의 방법론 추출	- 공식 인증 방법 - 적합성 평가를 위한 인증 프로그램 - 안전성 평가 검증 - 이해관계자 협업

	용어	측정	성능 특성화	호환성	규제 평가
AI 시스템 생성 및 생산 - 데이터 세트 및 알고리즘 (편향성 포함)/모델 - 사이버보안 - 시스템 설계 및 통합 - 업스케일링 및 평가 - 품질 관리		- 데이터 준비 및 유지 관리 - 새로운 프레임워크 및 하이브리드 표준화	- 스마트 안전 - 강화된 안전 센서 - 보조 주행 일관성, 센서, 테스트, 방법, 안전 점검, 신뢰성	- HRI 영향평가 - 표준 지침 식별	- 교육 인력의 학습 및 인증 - 센서를 포함한 시스템 수준 인증
데이터 생성 - 컴파일링, 준비, 편향성 검사 - 분석, 가공, 라벨링 - 라이선스 및 제한사항 - 공유 및 마케팅	- 하이브리드 표준화 - 데이터 준비 및 유지관리 - 표준 지침 식별	- 편향되지 않은 데이터 - 안전 지표 - 데이터 삭제	- 기준 및 기준치 식별 - 복잡하고 비구조적인 환경 - 맥락의 종속성	- 견고성을 위한 공식 방법 - 인간 상호 작용	

20) human-robot interaction : 인간-로봇 상호작용의 영향

PART Ⅴ 향후 논의 내용

- 본 보고서는 표준화를 위해 포용적이고 편향되지 않으며 신뢰할 수 있는 AI에 대한 데이터 품질 요구사항을 검토함
  - 먼저, 연구 및 산업에서 데이터 품질 평가 및 편향성 완화를 위한 수평적 이니셔티브를 검토하여 AI용 데이터세트의 생성 및 문서화를 평가한 다음, 이미 구축된 AI 모델에 대한 데이터 품질, 편향성 검사 및 완화를 검토함
  - 두 번째 단계에서는 6개의 특정 산업에서 AI 모델의 데이터 품질 요구사항과 잠재적인 표준화 요구사항에 대해 논의함. 해당 산업은 소비자와 업계의 관심이 가장 높은 분야이며, 그 중 일부 분야는 특정 사용 상황에서 위험이 높을 수 있으므로 시장에 출시되기 전에 엄격한 법적 의무를 준수해야 함: 1) 교육 및 고용, 2) 법 집행 및 공공, 3) 의료 및 헬스케어, 4) 산업 자동화 및 로봇 공학, 5) 금융, 6) 소셜 미디어, 콘텐츠 중재, 추천 시스템을 포함한 미디어

1 교육 및 고용

- 교육 및 고용 분야에서는 특히 데이터 품질 향상과 표준화가 시급함
  - AI 모델에서 두 가지 주요 과제가 확인되었음
    - AI 데이터가 어떻게 사용되는지의 투명성과, AI를 통해 발생할 수 있는 권력의 불균형과 불균형의 고착화를 막는 방법임
  - 본 워크숍에서는 알고리즘 구조화의 위험을 해결해야 한다는 필요성에 공감했으며, 해결책으로 사회적 대화 메커니즘, 데이터 상호주의 및 데이터 검증이 제시됨
  - 표준화를 달성하기 위해서는 측정 및 용어 정의가 더욱 명확히 이루어져야 함
    - 아직까지 공통 용어를 사용한 '편견 대 차별', '투명성 및 정확성'의 정의가 마련되지 못했으며, 교육 및 고용 분야에서 성과 평가에 대해 AI 모델이 수행할 일과 수행하지 말아야 할 일이 명확히 구분되지 않고 있음
    - AI 법은 AI 시스템을 규제하기 위한 위험 기반 접근 방식의 좋은 예시이며, 동시에 많은 딥 러닝 알고리즘의 블랙 박스 특성을 해결하기 위한 맞춤형 설명이 필요함

2 법 집행 및 공공 부문

- 법 집행 및 공공 분야에서는 얼굴 인식 기술뿐만 아니라 다른 AI 도구 및 응용 프로그램에도 데이터 품질 향상과 표준화가 시급함
  - AI에 의해 생성된 산출물(신원 확인된 사람의 성명, 차량 번호, 사이버 범죄 탐지 등)으로 인해 중요한 법적 판단(체포, 법원 판결 등)이 달라질 수 있기 때문에 많은 시민들이 법 집행 과정에서 AI를 활용하는 것에 대해 부정적 인식을 보임
    - 따라서 법 집행 및 공공 분야에서는 특히 신뢰성, 투명성 및 개인정보 보호 문제에 특히 주의를 기울여야 함
  - 이를 위해 해결해야 할 시급한 과제로 3가지가 제시됨
    - 첫째는 성별, 나이, 장애 및 인종에 따라 시민을 차별하지 않도록 보장하기 위한 메트릭 및 평가 규칙을 기반으로 정의할 수 있는 AI 시스템의 개선, 둘째는 법 집행 기관 간 통신 및 법 집행 기관과 비집행 기관의 통신의 표준화를 통한 상호 운용성 증진, 셋째는 법 집행 기관의 인력에 제대로된 AI 교육을 시행하는 것임

3 금융

- 금융 분야의 경우 데이터의 부적절한 사용으로 발생할 수 있는 문제 해결이 가장 중요함
  - 머신러닝과 같은 대용량 데이터 모델의 품질과 빅데이터의 품질 모델을 평가할 수 있는 평가 기준이 필요하며, 금융 분야 내 공통 용어 정의가 필요함
    - AI 시스템에서 편향성이 높은 결과를 도출할 경우 이에 따른 평가를 수행해야 하며, 표준 지침을 마련하여 편향성을 최소화해야 함

4 소셜 미디어, 콘텐츠 조정, 추천 시스템을 포함한 미디어

- 소셜 미디어, 콘텐츠 조정, 추천 시스템을 포함한 미디어 분야의 경우 편향성에 대한 정의를 합의하지 않았기 때문에 공통 용어를 정의하고 데이터세트를 감사하기 위한 합의된 공통 지침 및 표준을 제시해야 함
  - 미디어 분야는 텍스트, 비디오, 이미지 및 오디오를 포함하는 정보를 많이 다루는 분야로서, 다른 부문에 비해 개인 데이터를 많이 활용하기 때문에 개인 데이터에 대한 접근 논의가 많이 이루어짐
    - 이 때문에 개방성 제고가 중요하며, 미디어 분야에서 AI 모델이 어떻게 생성되었는지 정보를 연구 기관에 제공해 품질 보증을 받아야 함

- AI 모델의 품질을 테스트하고, 모델이 수행하는 작업의 적용 가능성을 비교할 수 있음

5 의료 및 헬스케어

- 의료 및 헬스케어 분야의 경우, AI 데이터 품질 요구사항과 관련하여 유럽 일반 개인정보보호법과 같은 많은 규제에 직면해 있음
  - 의료 및 헬스케어 분야는 설명 가능성과 투명성 제고를 위해 '블랙박스' 접근 방식에서 전환할 필요가 있음
  - 데이터 표준화를 위해서 좋은 데이터세트를 정의하는 방법, 데이터 시각화 방법, 알고리즘 편향성 극복 방법 등을 찾아야 함
    - 차별 없는 데이터세트를 통계에 반영하기 위해 표준화 전단계 연구가 필요하며, 재결합 가능한 데이터 측정지표를 설정하는 방법이 필요함
  - 다른 분야와 마찬가지로, 지침 제공 및 지침 부문 개발이 이미 확인됨

6 산업 자동화 및 로봇 공학

- 산업 자동화 및 로봇 공학 분야의 경우 모델 기반 인증 및 데이터 기반 인증 검증을 달성하는 것이 가장 중요함
  - 산업 자동화 및 로봇 공학 분야는 안전성과 시스템 견고성을 가장 중요한 측면 중 하나로 포함하며, 이를 위해 기술 지침이 필요함
    - 표준화 산출물이 광범위하게 활용되도록 하기 위해 제품 개발의 마지막 단계에서는 사용자를 고려해야 하며, 이 때문에 기술 지침이 필요함
  - 데이터 품질 요구사항에서 표준화가 필요한 부분은 스마트 안전, 인간의 기본권 보장 및 특수 클린 데이터 수집(special clean data collection)임
- 이처럼 AI는 다양한 부문에서 여러 수준으로 활용되고 있으며, 포용적이고 편향되지 않으며 신뢰할 수 있는 AI 시스템에 대한 품질을 보장하기 위한 AI법은 수평적으로 설계됨
  - AI법에서는 리스크를 감수하는 접근 방식을 선택했으며, 생명 및 건강에 대한 위험이 높을수록 품질 및 데이터 투명성 요구사항도 고도화되어야 함
  - 수평적 접근 방식이 일반적이나, 보편적 접근 방식을 보완하기 위해 부문별 표준도 필요할 것으로 보임
    - 표준화를 위한 용어, 분류 체계 및 교육 방식이 정립되고 있음

표준화를 진행시키기 위해 중소기업, 입법자, 소비자, 기본권, 기초과학적 관점에서 부문별 정보를 어떻게 결합시킬지 논의를 진전시켜야 함

- 정책적 관점에서 AI법이 당면한 주요 과제는 2025년 적용 시점임
  - 따라서, AI법이 적용되기 전까지 사용 가능한 표준이 필요하고 표준 개발 기관에 대한 권한을 부여해야 함
- 반면 AI 데이터 품질 요구 사항은 기본권, 윤리적 의미 및 다양한 이해관계자의 포괄적 참여와 같은 사회적 문제를 다루고 있음
  - 따라서 표준화 과정에 완전한 포용성을 보장하고 사회의 관점을 통합하도록 보장해야 함
- 또한, 향후 3년 동안 작업을 계획하기 위한 현 상황 격차를 식별해야 함

산업 및 정부 부문에서도 AI를 활용하기 때문에 고품질의 데이터세트는 정부 부문에서도 중요함

- 신규 표준화 전략(The New Standardisation Strategy)은 ESO 간의 조정, ESO에 구체적인 책임을 부여하는 명확한 로드맵, 자원 문제 해결, 유럽 연합 가치를 대변하는 계획 등 표준화 시스템의 효율성을 높이기 위한 여러 노력을 포함함
- 또한, 수평 표준화와 수직 표준화 사이에는 상호작용이 있어야 함
  - 일반 표준은 AI를 위한 기본 원칙을 수립하는 동시에, 부문 표준을 계획하고 평가함
- 앞으로도 더 나은 표준화 프로세스를 조정하기 위한 기회가 나타날 것으로 예상됨
  - 연구개발과 표준화 사이의 관계는 정치적 의제에서 상대적으로 중요하게 여겨지고 있으며, 유럽 표준화 기구 내에서 유럽 참가국들 사이에 조정이 더 필요할 것으로 여겨짐
- 모든 유럽 국가 표준화 기구에서 표준화 과정에 참여하고 있기 때문에 표준화가 원활히 이루어질 것으로 기대됨

PART Ⅵ

결론

본 보고서는 현재 AI 시스템의 과제와 편향성 해결을 위한 해결 방안을 제시함

- AI 시스템은 의료 서비스, 노동 조건, 생산성 향상뿐만 아니라 전반적인 경제 및 사회복지, 삶의 질을 개선할 수 있음
- 그러나, 동시에 최첨단 AI 시스템의 높은 편향성으로 인해 사회적으로 부정적인 영향을 미칠 수 있어 우려가 커지고 있음
  - 연구기관에서는 입력 데이터부터 시작하여 AI 모델 학습을 큐레이션함으로써 편향성을 완화할 수 있는 방법을 집중적으로 연구하고 있음
  - 그럼에도 불구하고 데이터 품질을 측정하는 기준이 아직 제대로 정립되지 않아 이에 대한 해결책이 필요한 상황임
- PSIS 워크숍의 목표는 기존 표준화 노력과 누락된 표준화 노력을 매핑하고, 새로운 표준 초안 작성 단계를 권장함
  - 워크숍에 모인 AI 실무자와 연구자들은 현재 AI에 존재하는 편향성을 평가하고 완화하기 위해 어떤 방법이 개발되었는지를 설명했음
  - 이를 통한 궁극적인 목표는 현재와 미래의 표준을 사용하여 AI 모델의 편향을 완화하는 방법에 대한 논의를 촉진시키는 것임
- AI 시스템 표준화를 위한 노력이 본격화되고 있는 가운데, AI 분야의 표준은 연구에서 시장으로의 기술 이전을 촉진할 수 있으며, 무엇보다도 유럽 표준이 가지는 매우 중요한 이점은 유럽의 윤리와 가치를 뒷받침하는 법률 시행을 지원하는 통일되고 합의된 요구사항을 수립할 수 있는 점임
- 본 워크숍은 상향식 접근 방식(전문가, 학계, 연구원, 기업, 이해관계자)과 하향식 접근 방식(공공 정책 및 법률)을 통합하여 공동 목표를 향해 목표를 제시했다는 점에서 의의가 있음
- 각 산업별로 특성이 다르기 때문에 특정 산업 및 응용 분야에 따라 각각의 표준이 필요할 수 있음. 이를 위해 수평적 표준과 수직적 표준 간의 중복 및 불일치를 방지하기 위해 해당 부문의 전문가 및 전문 기관과의 협력이 필요함
- 본 워크숍에서는 AI를 위한 데이터 표준을 정립하기 위해 아래와 같은 사항을 정리하고, 향후 유럽연합의 기준을 지원하기 위해 노력을 지속하기로 결정함

[AI 표준화 로드맵에서 다룰 수 있는 표준화 기회와 요구사항]

AI 시스템 개발 표준

- 관리 표준(데이터 품질, 시스템 보증, 편향성 배제)
- 배포 표준(이해, 신뢰, 일관된 적용)
- 운영시험 표준(배포 전후 구축)
- 데이터 품질 측정 표준(지표, KPI)
- AI 알고리즘 인증 표준

AI 시스템 배포 표준

- AI, 데이터, 시스템, 용어, 위험, 한계, 안전, 인간/AI 인터페이스, 데이터 품질의 정의
- 교육 및 학습(AI의 의미론, 의도, AI 시스템 설계 및 계획, 데이터 요구사항, 편향성 측정 및 완화를 위한 적용 및 접근 방식 등)
- AI 시스템 감독 표준(당국별) 및 배포 표준(사용자와 업계간, 중소기업 포함) 설계
- AI를 위한 데이터 생성(데이터의 입증, 수집 및 발표, 데이터 사용 및 공유, 데이터 세트 적합성 체크리스트, 신뢰할 수 있는 데이터, 데이터 유지 및 견고성)

산업별 인공지능 활용을 지원하는 표준

- 전 산업의 상용 AI 모델 품질 보장
- 중소기업(모든 유형의 조직 및 단체가 AI 시스템에 접근할 수 있는 표준)
- 법 집행(게이트 키핑, 안면 인식, 이미지 품질 등)
- 금융(데이터 '공정성', AI 기반 의사결정 과정, 객관성 및 정량적 측정)
- 미디어(미디어용 AI 편향성의 구체적인 측면, 다양한 AI 모델 비교, 데이터 세트 벤치마킹, 미디어에 특화된 편향성 및 품질 정의, 기존 모범 사례 체계화, 데이터의 편향성 테스트 및 벤치마킹)
- 의료(규제 상황(GPDR, MDR), 규제 당국의 데이터 관련 조항, 법률 준수, 다양한 사용 사례, 공통 표준 용어 정립, 데이터의 익명화, AI 시스템 문서화, 개인화 요구사항)
- 산업 자동화 및 로보틱스(AI 시스템의 안전(설계에 의한 안전성) 및 지표, 다양한 응용 분야, 로봇/인간 인터페이스 및 안전, 인권 존중, 프라이버시), 용어(사용자가 표준화에 기여하는 데 도움을 줄 것임).

참 고 문 헌

- Balahur, Alexandra; Jenet, Andreas; Hupont Torres, Isabelle; Charisi, Vasiliki; Ganesh , Ashok; Griesinger, Claudius B.; Maurer, Philip; Mian, Livia; Salvi, Maurizio; Scalzo, Salvatore; Sol er G arri do, Jose p; Taucer, Fabio; Tolan, Songül, 「Data quality requirements for inclusive, non-biased and trustworthy AI. Putting-Science-Into-Standards」, Publications Office of the European Union, Luxembourg, (2022)
- European Commission., 「European Commission's proposal for a Regulation on Artificial Intelligence」, Luxembourg: European Commission, (2021)
- Gebru, T, J Morgenstern, B Vecchione, J Wortman Vaughan, H Wallach, H Daume III, and K Crawford, 「Datasheets for Datasets」, https://arxiv.org/abs/1803.09010., (2018)
- Hupont, Isabelle, and M Chetouani, 「Region-based facial representation for real-time action units intensity detection across datasets」, Pattern Analysis and Applications 22 (2): 477-489, (2019)
- Joint Research Centre (European Commission), 「Data Quality Requirements for Inclusive, Non-Biased and Trustworthy AI」, https://op.europa.eu/en/publication-detail/-/publication/b11a0504-75eb-11ed-9887-01aa75ed71a1/language-en/format-PDF/source-291783199, (2022)
- Sun, Tony, Andrew Gaut, Shirlyn Tang, Mai ElSherief, Yuxin Huang, Diba Mirz, and Jieyu Zhao. 「Mitigating Gender Bias in Natural Language Processing: Lit erature Review」 Association for Computational Linguistics, (2019)
- Tolan, S, A Pesole, F Martínez-Plumed, E Fernández-Macías, J Hernández- Orallo, and E Gómez., 「Measuring the occupational impact of AI: tasks, cognitive abilit ies and AI benchmarks」, Journal of Artificial Intelligence Research, 71, 191-236., (2021)
- World Health Organisation, 「Ethics and governance of artificial intelligence for health」, https://www.who.int/publications/i/item/9789240029200, (2022)

데이터산업 동향 이슈 브리프

ISSUE BRIEF

| **발행일** 2023년 9월 27일

| **발행처** K data 한국데이터산업진흥원
서울시 중구 세종대로 9길42, 부영빌딩 7, 8, 11층

| **기획 및 편집** 산업기반본부 산업기획팀 하진희 팀장,
이정현 연구위원, 나현대 책임

| **문의처** Tel: 02-3708-5362