

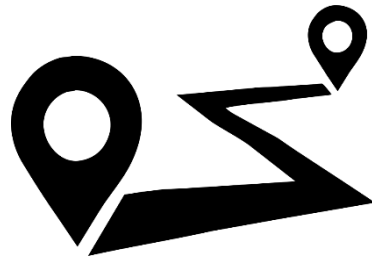
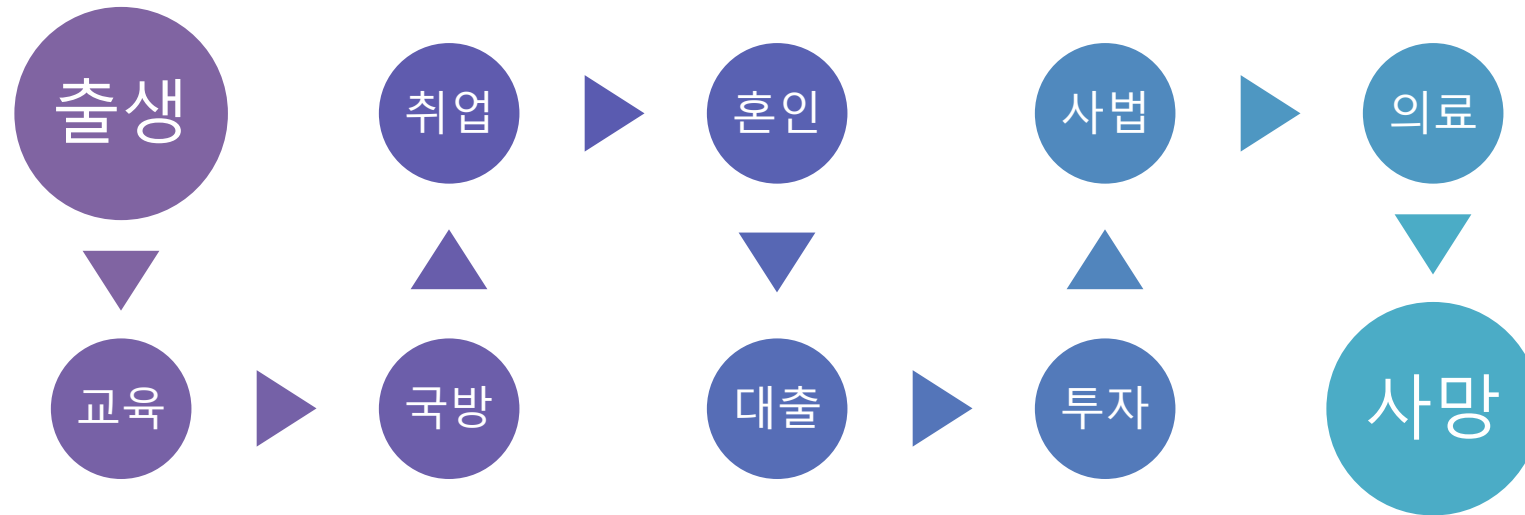
다가온 AI 규제 시대, 어떻게 AI 비즈니스를 준비할 것인가?

2021. 12. 9.

Graduate School of Innovation and Technology Management,
KAIST

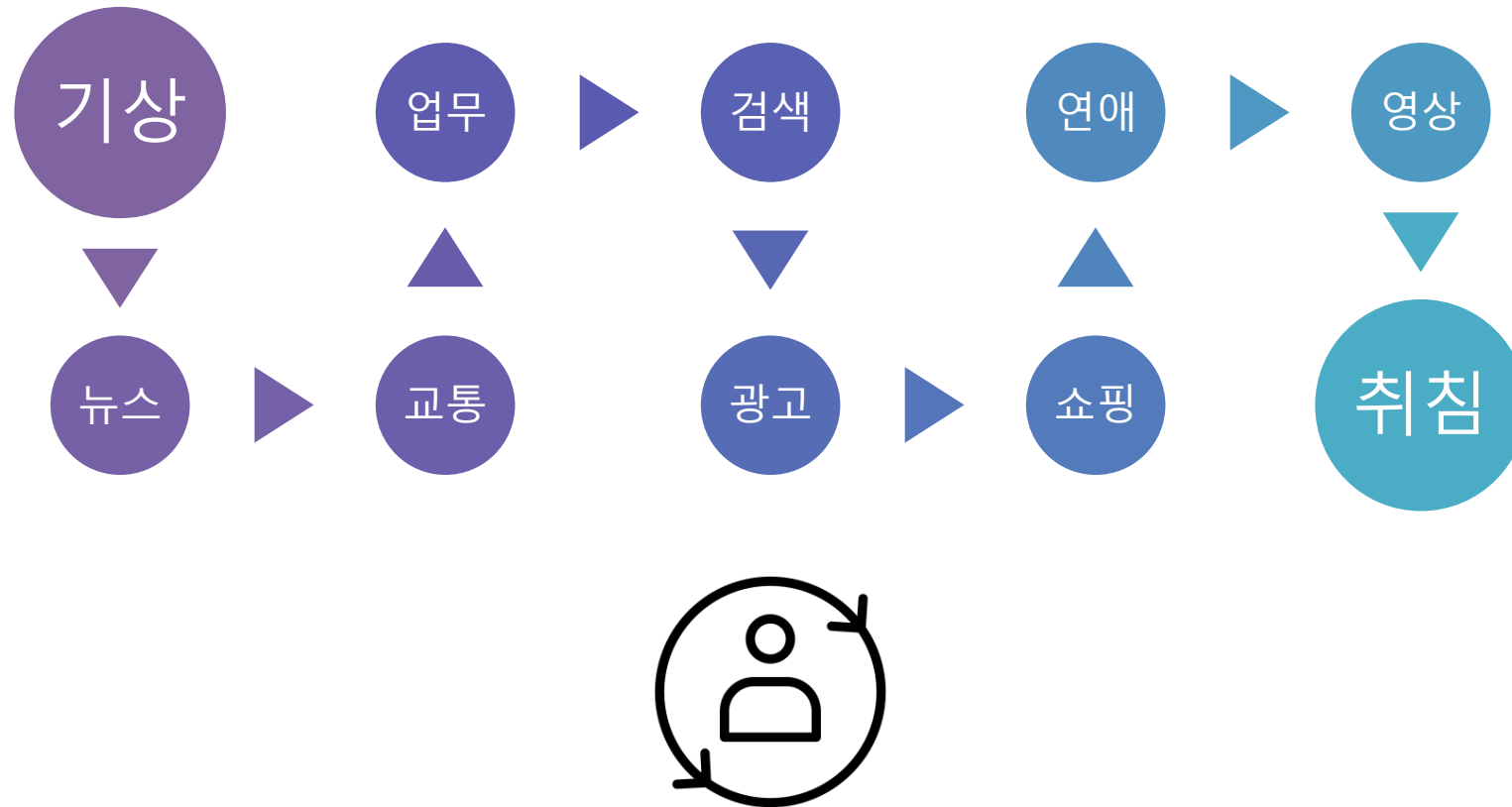
Byoungpil Kim

우리 삶의 중요한 순간들



→ AI가 우리의 삶의 궤적을 결정

우리의 일상



➔ 우리 일상의 매 순간마다 AI가 관여

(전제) 인공지능 = 함수

$$\underline{f}(\underline{\quad}) = \underline{\quad}$$

인공지능 입력 변수 출력 변수

인공지능의 활용 사례

$$f(\text{ waveform }) = \text{"Siri야"}$$

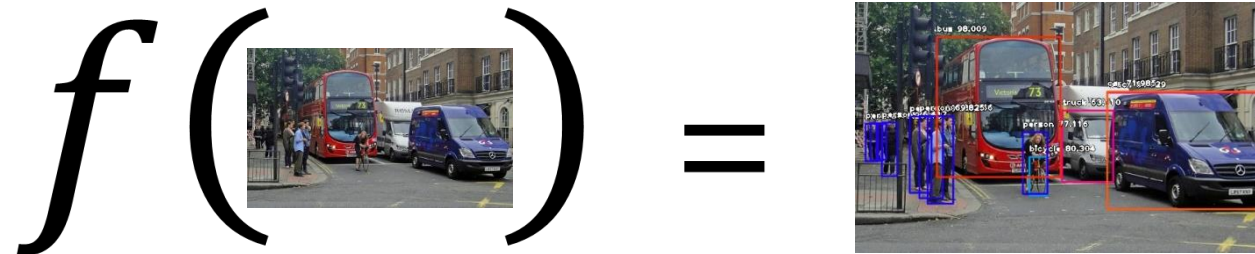
$$f(\text{ 你好 }) = \text{"안녕하세요"}$$

$$f(\text{ chest_xray_image }) = \text{폐렴이 있는지?}$$

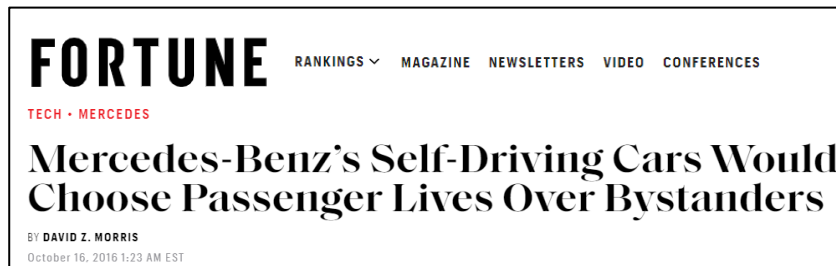


과연 어떤 위험이 있을까?

인공지능의 위험 (1) 생명권·신체의 자유



의료 / 자율주행 / 로봇틱스 / 사법 AI 등
→ AI 안전성 검증 시스템 필요



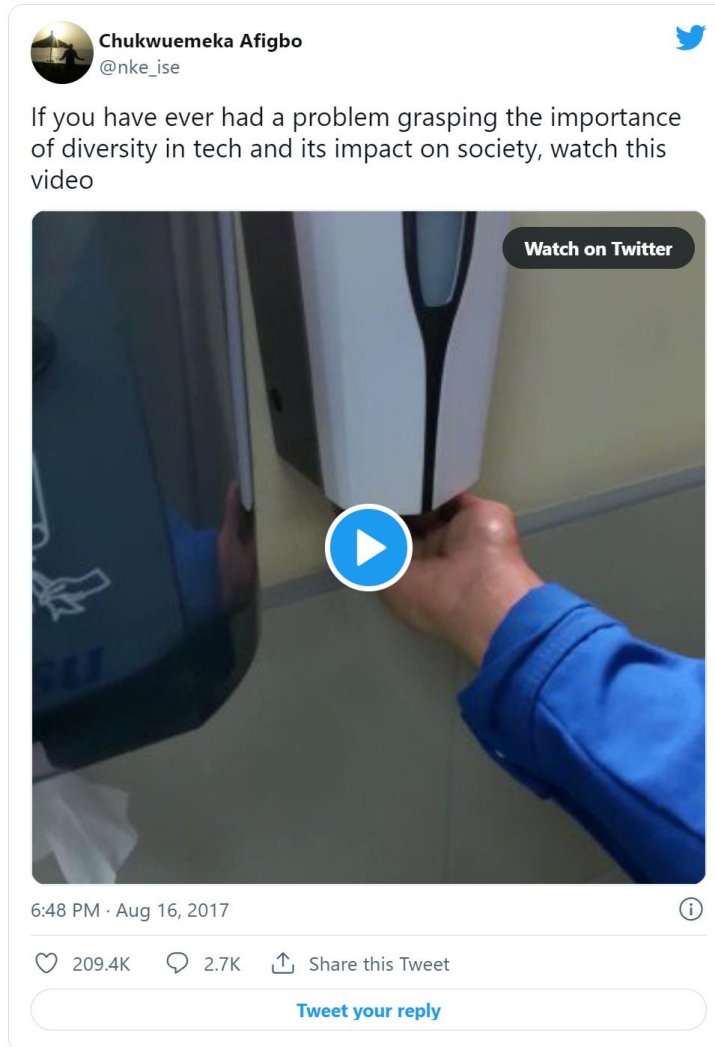
인공지능의 위험 (2) 기회의 균등

$$f(\text{Resume of John Doe}) = \text{자격이 있는지?}$$

취업 / 교육 AI 등
→ AI 차별 검증 시스템 필요



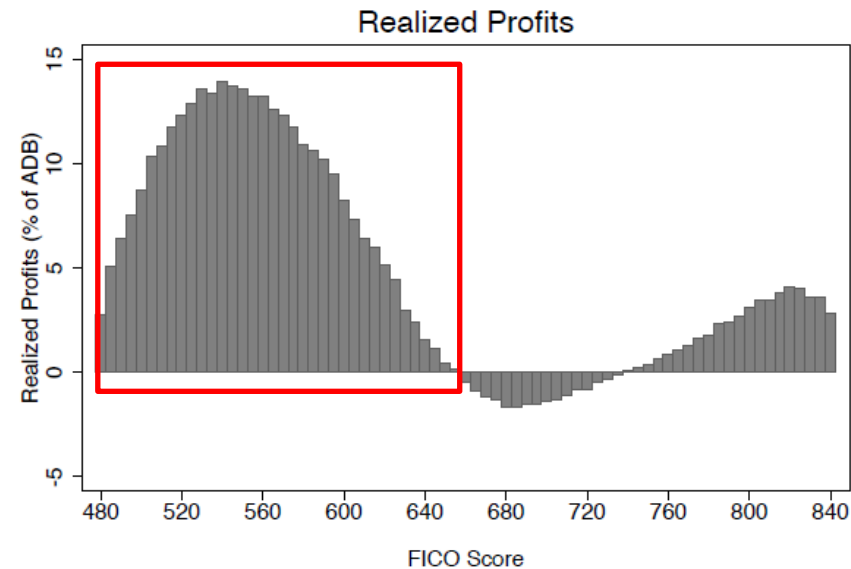
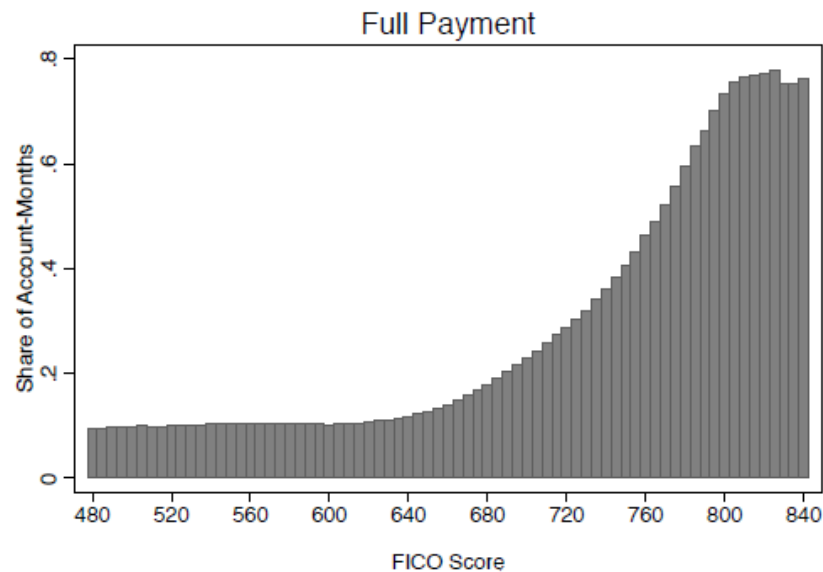
Racist Soap Dispenser



인공지능의 위험 (3) 소득의 분배

$$f\left(\text{Icon of hands exchanging money}\right) = \text{신용도?}$$

금융 / 상품 추천 / 광고 AI 등
→ AI에 의한 소득 재분배 효과?



교과서 가격이 \$2,198,177.95?



The Making of a Fly: The Genetics of Animal Design (Paperback) by Peter A. Lawrence

[Return to product information](#)

Always pay through Amazon.com's Shopping Cart or 1-Click.
Learn more about [Safe Online Shopping](#) and our [safe buying guarantee](#).

Price at a Glance

List Price: \$70.00

Used: from **\$35.54**

New: from **\$1,730,045.91**

Have one to sell? [Sell yours here](#)

All

New (2 from \$1,730,045.91)

Used (15 from \$35.54)

Show ☒ New ☐ Prime offers only (0)

Sorted by Price + Shipping

New 1-2 of 2 offers

Price + Shipping	Condition	Seller Information	Buying Options
\$1,730,045.91 + \$3.99 shipping	New	<p>Seller: profnath</p> <p>Seller Rating: ★★★★★ 93% positive over the past 12 months. (8,193 total ratings)</p> <p>In Stock. Ships from NJ, United States. Domestic shipping rates and return policy.</p> <p>Brand new, Perfect condition, Satisfaction Guaranteed.</p>	<p>Add to Cart</p> <p>or</p> <p>Sign in to turn on 1-Click ordering.</p>
\$2,198,177.95 + \$3.99 shipping	New	<p>Seller: bordeebook</p> <p>Seller Rating: ★★★★★ 93% positive over the past 12 months. (125,891 total ratings)</p> <p>In Stock. Ships from United States. Domestic shipping rates and return policy.</p> <p>New item in excellent condition. Not used. May be a publisher overstock or have slight shelf wear. Satisfaction guaranteed!</p>	<p>Add to Cart</p> <p>or</p> <p>Sign in to turn on 1-Click ordering.</p>

금융 기회와 차별

세계일보

대출하러 은행갔더니 "고졸? 이자 더 내라"

입력 : 2012-07-23 19:24:59 | 수정 : 2012-07-24 08:48:52

<http://www.segye.com/newsView/20120723023137>

The New York Times

***Apple Card Investigated After
Gender Discrimination Complaints***

A prominent software developer said on Twitter that the credit card was “sexist” against women applying for credit.

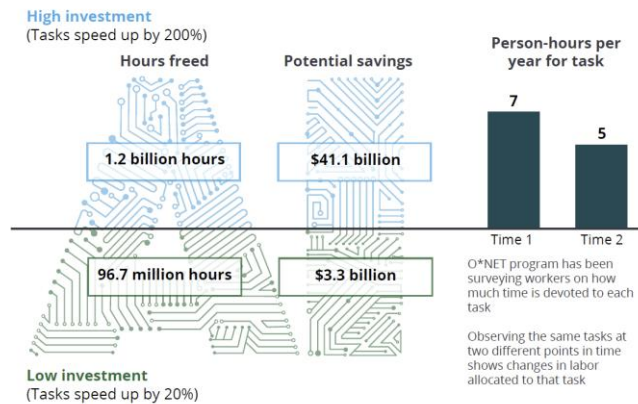
<https://www.nytimes.com/2019/11/10/business/Apple-credit-card-investigation.html>

인공지능의 위험 (4) 적법절차 원칙

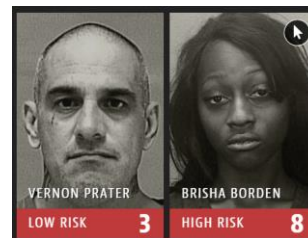
$$f(\text{Surveillance Footage}) = \text{범죄 가능성?}$$

행정 / 경찰 / 사법 / 플랫폼 규율 AI 등
→ AI가 적법절차 원칙 준수?

Figure 1. How much savings can AI in government generate?



Source: Deloitte analysis.



Banned by Amazon for returning faulty goods

Greg Nelson has bought 343 items from the online giant since 2014. But after sending 37 back he is now blocked from using it and can't reclaim his credit balance

사법 AI

- COMPAS - 미국 사법부의 재범 예측 인공지능 활용

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica
May 23, 2016

Two Petty Theft Arrests



VERNON PRATER

LOW RISK 3



BRISHA BORDEN

HIGH RISK 8

Borden was rated high risk for future crime after she and a friend took a kid's bike and scooter that were sitting outside. She did not reoffend.

Two Petty Theft Arrests

VERNON PRATER

Prior Offenses
2 armed robberies, 1 attempted armed robbery

Subsequent Offenses
1 grand theft

LOW RISK 3

BRISHA BORDEN

Prior Offenses
4 juvenile misdemeanors

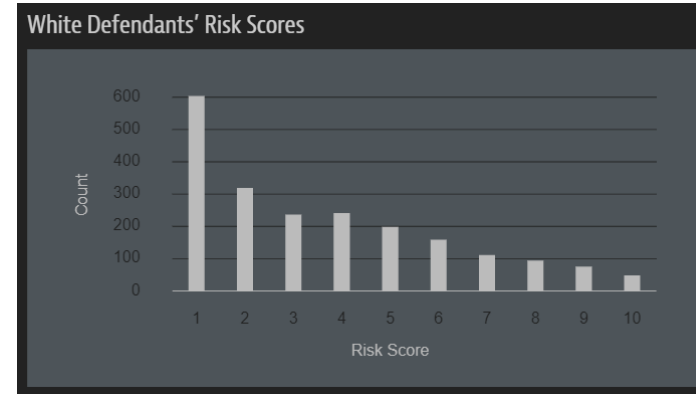
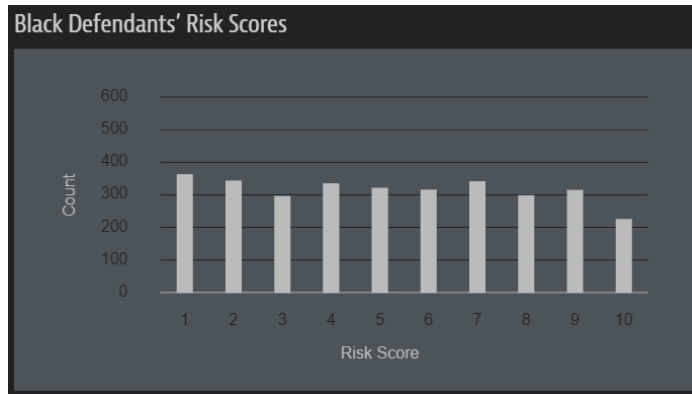
Subsequent Offenses
None

HIGH RISK 8

Borden was rated high risk for future crime after she and a friend took a kid's bike and scooter that were sitting outside. She did not reoffend.

제기된 문제점

- 인종별 위험도 분포 차이



- 인종별 오류율 차이

Prediction Fails Differently for Black Defendants		
	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

Overall, Northpointe's assessment tool correctly predicts recidivism 61 percent of the time. But blacks are almost twice as likely as whites to be labeled a higher risk but not actually re-offend. It makes the opposite mistake among whites: They are much more likely than blacks to be labeled lower risk but go on to commit other crimes. (Source: ProPublica analysis of data from Broward County, Fla.)

<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

인공지능의 위험 (5) 표현의 자유, 민주주의

$$f(\text{icon}) = \text{유해 콘텐츠?}$$

콘텐츠 추천 / 심사 AI 등
→ 표현의 자유 보장 vs 가짜 뉴스 막기

YouTube will rely more on AI moderation while human reviewers can't come to the office

Mistakes will be made

By Jacob Kastrenakes | @jake_k | Mar 16, 2020, 3:06pm EDT

Facebook is flagging some coronavirus news posts as spam

As social distancing limits in-office staff, social platforms are relying more on AI to moderate their content.

By Rebecca Heilweil | Updated Mar 17, 2020, 9:45pm EDT

How to Burst the "Filter Bubble" that Protects Us from Opposing Views

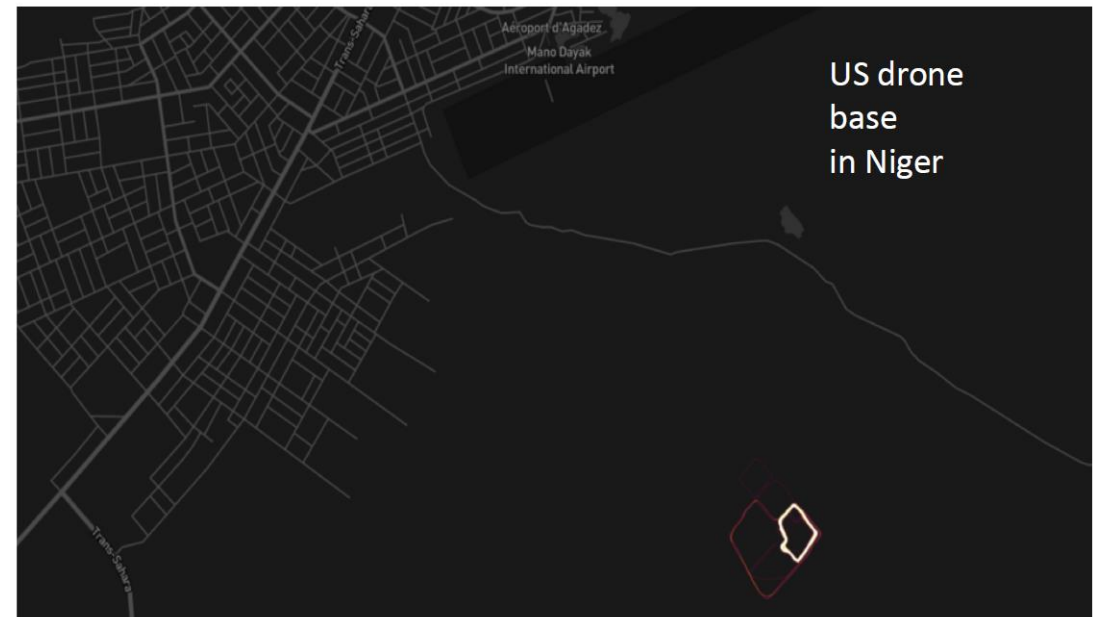
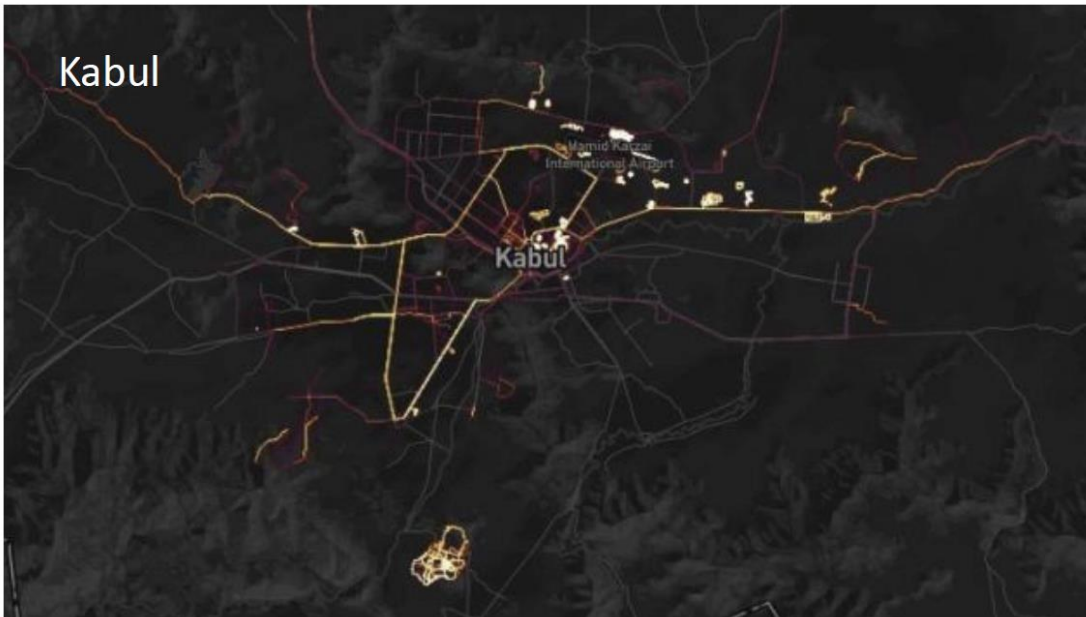
Computer scientists have discovered a way to number-crunch an individual's own preferences to recommend content from others with opposing views. The goal? To burst the "filter bubble" that surrounds us with people we like and content that we agree with.

by **Emerging Technology from the arXiv**

November 29, 2013

국가안보에의 위협

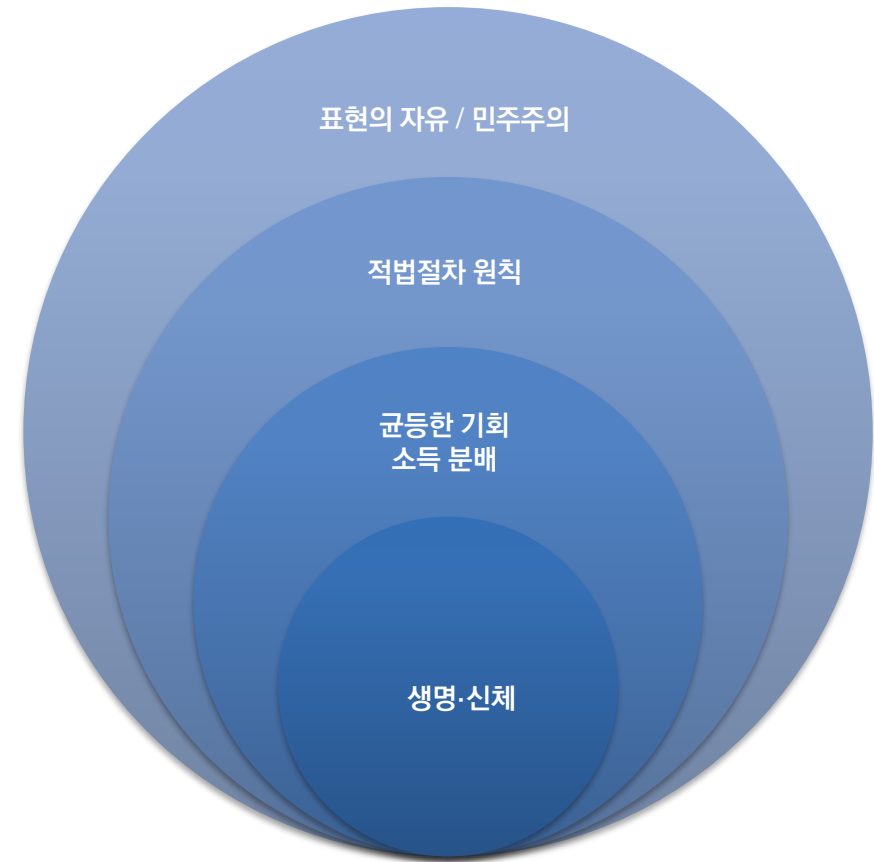
- 자율살상무기체계의 등장, 사이버 안보 위협의 증가
- 빅데이터 분석을 통한 국가안보 위협 사례
 - 미국 군사기지 위치 유출



우리가 공정한 사회에 살 수 있을 것인가?

→ 인공지능에 의해 결정될 것임

$f()$



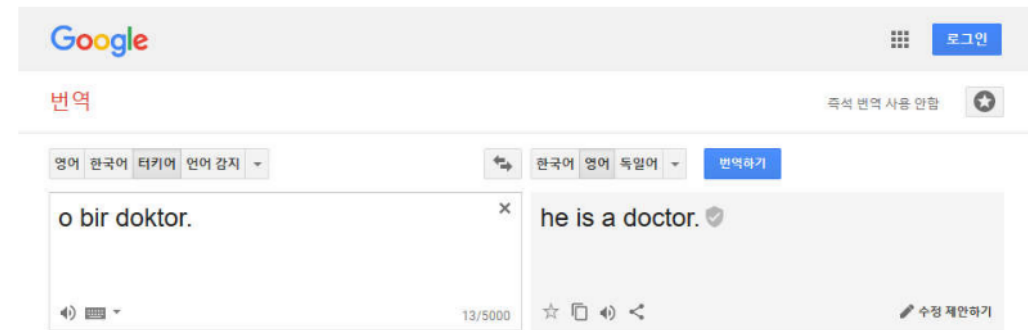
집단 대표의 문제

- 검색 엔진, 기계 번역, 언어모델 등
- 학습 데이터에서 대표 (Representation) 되지 못하는 소수자 집단에 대한 배제/차별

→ 추가적 인간 개입(intervention)이 필요함!



자료 4. 번역기에 의한 터키어 'o bir doktor.'의 영어 번역 결과



자료 7. 번역기에 의한 터키어 'o bir hemşire.'의 영어 번역 결과



수정된 “CEO” 검색 결과



Marriott CEO Arne Sorenson Is Th...
chiefexecutive.net



Chief executive officer - Wikipedia
en.wikipedia.org



Proximus approves CEO appointment -...
mobileworldlive.com



Casey's Announces C...
us.acrofan.com



McDonald's C.E.O. Fired Over a Relatio...
nytimes.com



ESSENCORE
essencore.com



How to use 'CEO magic' when tryin...
europeanceo.com



Rise of the next-gen bank CEO
americanbanker.com



Harvard study: What CEOs do all day
cnbc.com



HSBC's Caretaker CEO Sets Sight...
bloomberg.com



CEO Message | Company | AGC
agc.com



What do CEOs do? A CEO Job Description - ...
steverobbins.com



아우디, BMW 마르쿠스 듀스만(Markus Du...
aitimes.kr



Selective CEO begins the next chapter of his ...
insurancebusinessmag.com



TOP CEO 56] 전창원 빙그레 대표이사 - CEO ...
ceomagazine.co.kr



CEO | Company Introduction | C...
lgchem.com

번역 오류의 수정

언어 감지 터키어 일본어 영어 ▼ ↔ 한국어 영어 일본어 ▼

o bir doktor × 번역이 성별에 따라 달라집니다. 자세히 알아보기 ☆

he is a doctor (남성) 🔊 📄

She is a doctor (여성) 🔊 📄

12/5000 🔊 🔊

의견 보내기

언어 감지 터키어 일본어 영어 ▼ ↔ 한국어 영어 일본어 ▼

o bir hemsire × 번역이 성별에 따라 달라집니다. 자세히 알아보기 ☆

he is a nurse (남성) 🔊 📄

she is a nurse (여성) 🔊 📄

13/5000 🔊 🔊

의견 보내기



인공지능, 어떻게 규율해야 하나?

인공지능 윤리원칙의 “범람”

SAPI AI윤리 원칙 인덱스 (SAPI AI ETHICS & GOVERNANCE GUIDELINES REPOSITORY)

그래프 (면적 / 꺾은선 / 파이)

기준일: 2021.2.17. 현재

수집 범위: 준칙 형태로 되어 있고, 해외의 경우 영문/영역이 제공된 경우로 한정

총 81개의 검색 결과

네이버

네이버 AI 윤리 준칙

Republic of Korea | 2021. 02. | Private Sector

Fairness / Non-Discrimination / Inclusiveness | Transparency / Explainability | Privacy | Safety / Security / Reliability | Beneficence / No Harm

과학기술정보통신부/정보통신정책연구원

국가 인공지능 윤리기준

Republic of Korea | 2020. 12. | Government/Public Sector

Fairness / Non-Discrimination / Inclusiveness | Accountability / Responsibility / Auditability | Transparency / Explainability | Privacy | Safety / Security / Reliability | Beneficence / No Harm | Human Dignity / Autonomy

Ping An

AI Ethics Governance Statement and Policies of Ping An Group

China | 2020. 12. | Private Sector

Fairness / Non-Discrimination / Inclusiveness | Accountability / Responsibility / Auditability | Transparency / Explainability | Safety / Security / Reliability | Human Dignity / Autonomy | Human Intervention / Oversight

과기정통부, 사람이 중심이 되는
「인공지능(AI) 윤리기준」 마련

- 공개 공청회(12.7) 등 각계 전문가·시민 공개 의견수렴 거쳐 발표 -
- ‘인간성(Humanity)을 위한 인공지능(AI)’의 3대 원칙·10대 요건 담아 -

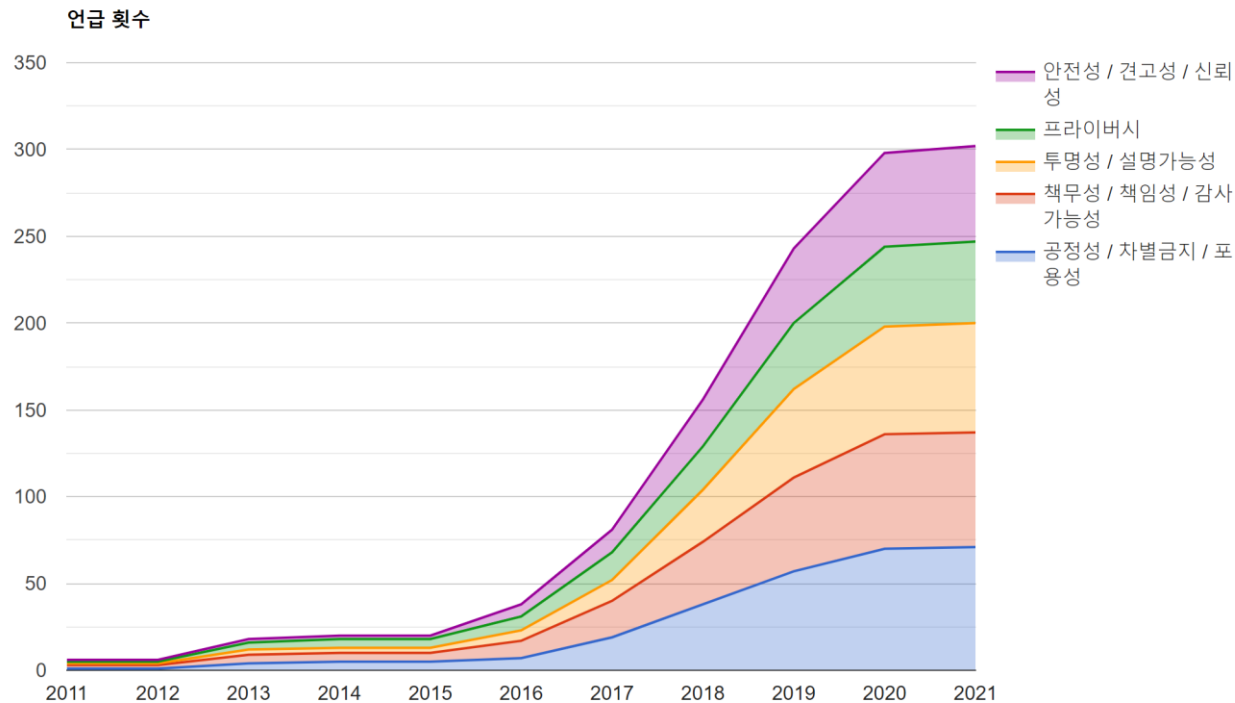


제 목 : 「금융분야 인공지능(AI) 가이드라인」이 시행됩니다.
- 금융권 AI 활용을 활성화하고 AI 기반 금융서비스에 대한 신뢰를 제고하기 위한 모범규준 마련·발표 -

인공지능 윤리원칙의 주요 내용

- 공정성 / 차별금지 / 포용성
- 책무성 / 책임성 / 감사가능성
- 투명성 / 설명가능성
- 프라이버시
- 안전성 / 견고성 / 신뢰성

➔ 왜 AI 윤리원칙이 범람하고 있을까?



<http://repository.sapi.co.kr/area-chart/>

Laws of Fear – 위험 인식

- **위험에 대한 과대 평가**
 - 대중은 사소한 위험에 대해 과도한 두려움을 가질 수 있음
 - 인간의 인지 한계 – 가용성 편향 등
 - 선별된 언론 보도 사례를 통해 위험을 인지
- **위험에 대한 과소 평가**
 - 대중은 중대한 위험에 대해 충분한 두려움을 갖지 못할 수 있음
 - 과학의 근원적 한계와 불확실성 – “확증”을 위한 상당한 시간 소요
 - 과학에 대한 불신 – “과학의 정치화(politicalization of science)”
- **어떻게 적절한 규제 지점을 찾아낼 수 있을 것인가?**
 - 사전적(ex ante) 규제 방식
 - 사후적(ex post) 규제와 손해배상 방식





EU AI 규제 법안 (2021. 4.)

Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS

EU AI 규제 법안 – 배경

- **전문가 의견 수렴**

- High-Level Expert Group on AI (HLEG) 수립
- 2019년 – “Ethics Guidelines for Trustworthy AI” 발간
- 2020년 – “Assessment List for Trustworthy Artificial Intelligence (ALTAI)” 발간

- **공중 의견 수렴**

- 1,215건 의견 수렴 – 기업 352건, 개인 406건, 학계 152건, 공공기관 73건
- **AI에 대한 조치가 필요하다는 점은 대부분 동의함** (→ 현재 입법 공백 상태가 존재)
- 반대로 중복 입법, 의무간 충돌 및 과잉 규제 우려도 제기
- 대부분 AI에 대한 명확하고, 정교한 정의 조항의 필요성 제기
 - “위험”, “고위험”, “저위험” “원격 생체인식” 및 “위해”에 대한 정의가 중요
- 대부분 위험기반 접근법을 명시적으로 선호
- 기업들은 규제 샌드박스를 선호
- 집행 방식 – 50% 이상의 의견이 (i) **사전적 위험 자기 평가**, (ii) 고위험 AI 시스템에 대한 **사후적 집행 방식**을 결합하는 형태를 선호

EU AI 규제 법안 – 개요

- **입법 목적**

- AI 시스템의 발전·활용과 비례적 위험기반 접근 사이의 균형을 달성
- 기본권 및 안전에 심대한 위험을 낳을 수 있는 경우에만 규제 부담 부과
- 고위험이 아닌 경우 제한적인 투명성 의무만 부과, 자율 규제 권장

- **핵심 규제 내용**

- “수용 불가능한 AI”를 규정하고 이를 금지
- “고위험 AI”를 정의하고 의무사항을 부과
 - AI 생애주기 전체에 걸쳐 안정성 확보를 위한 규제 적용
- “제한적 위험 AI”에 대해서는 최소한의 투명성 의무 부과
- “낮은 위험 AI”는 규제를 두지 않음

EU AI 규제 법안 – 고려된 입법 대안

- **고려된 입법 대안**

- 1안 – 사업자의 자발적 라벨링 규제
- 2안 – 분야별, “ad-hoc” 접근
- 3안 – 비례적 위험기반 접근에 따른 EU 차원 입법
- 3+안 – 비례적 위험기반 접근에 따른 EU 차원 입법 + 非고위험 AI에 대한 행동강령 방식
- 4안 – 모든 AI 시스템에 대한 EU 차원 의무 부과 입법

- **선택된 입법 방안 – 3+안**

- 의무 준수 비용 최소화 및 불필요한 가격 상승 방지
- 중소기업(SME) 불이익에 대한 대비 – 규제 샌드박스 및 SME의 이익을 위한 조치
- AI에 대한 사회적 신뢰 증진, 법적 안정성 확보

- **예상 규제 준수 비용**

- AI 공급자 – 2025년까지 6,000~7,000 유로 추가 부담
 - (170,000 유로 상당의 고위험 AI 시스템에 대한 추산)
- AI 이용자 – 인간 감독 비용 연간 5000~8000 유로 부담
- 고위험 AI 검증 비용 – 3,000~7,500 유로

(1) 수용 불가능한 AI 위험

- **금지되는 AI 활용**

- 무의식적으로 인간을 조작할 수 있는 중대한 잠재성이 있는 경우
- 아동·장애인 등 특정한 취약 계층의 취약성을 악용하여 피해를 가할 수 있는 경우
- 공공기관에 의한 AI 기반 사회 평점 시스템
- 법집행 목적의 공개 장소에서의 **실시간** 원격 생체인식 신원확인 시스템(사법 통제 대상)

- **법집행 목적의 공개 장소 실시간 원격 생체인식 신원확인 시스템에 대한 사법 통제**

- 아래의 경우에만 사법기관/독립행정기관의 사전 승인을 얻어 허용
 - 특정 잠재적 범죄 피해자, 미아 탐색
 - 자연인의 생명 또는 물리적 안전에 대한 구체, 중대, 현존 위협 또는 테러 공격의 예방
 - 장기 3년 이상의 범죄자의 탐지, 위치확인, 식별 및 기소
- 다음과 같은 사항을 고려
 - 해당 사용이 필요한 상황의 성격, 특히 해당 시스템을 활용하지 않을 경우 초래될 수 있는 위협의 심각성, 가능성 및 정도
 - 해당 시스템을 사용함으로써 발생할 수 있는 관련 개인의 권리와 자유에 대한 영향, 특히 이러한 결과의 심각성, 가능성 및 정도

(2) 고위험 AI 시스템

- **고위험 AI 시스템의 유형**
 - 안전과 관련된 AI 시스템
 - 고위험으로 추가 규정된 AI 시스템
- **안전과 관련된 AI 시스템(부속서 II에 기재된 EU 법이 적용되는 경우)**
 - AI 시스템이 안전 부품(AI 시스템 그 자체가 제품인 경우 포함)인 경우
 - 완구, 레저기구, 승강기, 폭발물, 고압기구, 케이블, 의료기기 등 관련 규제
 - 민간 항공 안전, 차량, 농업·임업용구, 해상기구, 철도 등 관련 규제 (제3자의 적합성 평가가 필요한 경우)
- **고위험으로 추가 규정된 AI 시스템(부속서 III에 기재 – 개정 가능)**
 - 생체인식 신원확인 및 자연인의 분류
 - 중요 인프라의 관리 및 운영
 - 교육 및 직업 훈련
 - 채용, 근로자 관리 및 자기고용에 대한 접근
 - 이른바 플랫폼 노동에서의 작업 할당 및 성과 감독 포함
 - 핵심적 민간 및 공공 서비스 및 혜택에 대한 접근 및 이용
 - 자연인에 대한 신용평가 포함 – 소규모 제공자가 스스로 사용하는 경우 제외
 - 법집행
 - 이민, 망명 및 국경 관리
 - 사법 및 민주적 절차

EU AI 법안상 고위험 AI 핵심 규제 = 적합성 평가

- **안전과 관련된 AI 시스템 → 제3자 평가(기존 CE 제품 적합성 평가 제도 적용)**
 - CE 마킹 제도 – 제조자/수입자가 EU 제품 규제 준수 여부를 평가하도록 강제
 - (제품의 성능/품질에 대한 인증이 아님)
 - 제3자에 의한 사전적 적합성 평가를 요구(기존의 관련 규제의 절차에 따름)
- **부속서 III에 규정된 AI 시스템 → 원칙적 자기 적합성 평가**
 - 공급자에 의한 내부 통제 확인
 - 사전적 준수 여부에 관한 완전하고, 효과적이며, 적절하게 문서화된 평가
 - EU 데이터베이스에 등록
 - 단, 원격 생체인식 신원확인 시스템은 제3자 적합성 평가 필요



고위험 AI – 의무사항

- **위험관리 시스템(제9조)**
 - 고위험 AI 시스템에 대한 위험관리 시스템을 설립, 구현, 문서화 및 유지 필요
 - AI 전체 생애주기에 걸친 지속적이고 반복적 절차, 주기적 업데이트 필요
- **고품질 데이터 – 데이터와 데이터 거버넌스(제10조)**
 - 학습, 검증 및 시험 데이터셋은 관련성, 대표성, 무오류성, 완전성을 갖추어야 함
- **기술 문서 작성 및 로그 기록 보관(제11조, 제12조)**
 - 시장 출시 전 법적 요구사항 준수를 보여주는 기술 문서 작성
 - AI 시스템이 작동하는 동안의 자동적 사건 기록(로그) 기능 구현 – AI 시스템 기능을 생애주기 전체에 걸쳐 추적할 수 있는 수준
- **투명성 및 이용자에 대한 정보 제공(제13조)**
 - 이용자가 시스템의 출력을 해석하고 적절하게 이해할 수 있도록 충분히 투명해야 함
 - AI 시스템이 이용자에게 제공해야 할 정보 - 공급자 정보, AI 시스템의 성능(목적, 정확도, 오용 위험성, 성능 및 데이터 정보), 변경 사항, 인간 감독, 예상 수명 등

고위험 AI – 의무사항 (계속)

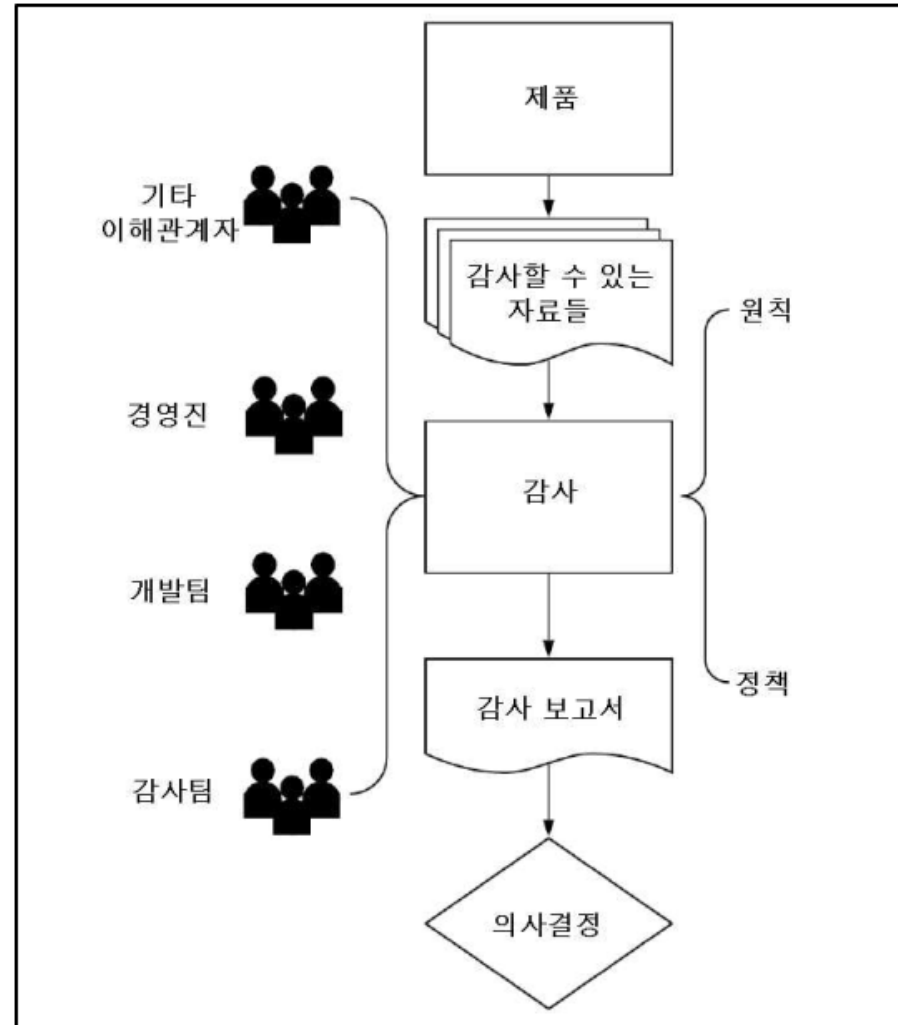
- **인간 감독(제14조)**

- 사용 기간 중 자연인에 의해 효과적으로 감독될 수 있도록 설계 및 개발(인터페이스 도구 포함)
- 정상 동작 또는 합리적으로 예견가능한 오용 상황에서 위험의 예방 또는 최소화
- 인간 감독이 AI 시스템에 내재화되거나 이용자에 의해 구현
- 이용자가 AI 시스템의 능력과 한계를 이해, AI 시스템에 의한 편향의 자동화 가능성을 인지, 결과물을 해석 가능, 인간이 반복 가능, “stop” 버튼 포함

- **정확성, 강건성 및 사이버 보안(제15조)**

- 정확성 – 목적에 비추어 적절한 수준의 정확성 확보
- 강건성 – 시스템이나 외부 환경 오류에 대한 저항성, ‘**피드백 루프**’에 대한 적절한 저감 조치
- 사이버 보안 – 외부 공격에 대한 저항성 및 적절한 기술적 조치 도입
 - 데이터 오염공격(data poisoning), 적대적 사례(adversarial examples)에 대한 기술적 조치 포함

AI 위험관리 시스템 – 내부 감사 필요!










어떻게 준수하나? AI 신뢰성 "기술표준"

기술표준 수립 동향

- AI 신뢰성 확보를 위한 국제적 기술표준 제정 시도가 활발히 진행 중
 - ISO/IEC 공동위원회 - AI 신뢰성 관련 기술 표준 다수
 - IEEE - 자율지능시스템 윤리 표준 다수
- 국내 TTA - 인공지능 신뢰성 개발 가이드북
- ➔ 대부분 현재 제정 작업이 진행 중인 상태

국제 기술표준의 실질적 규범화

- 향후 EU AI 규제 법안의 실제 입법화에 맞추어 세부적인 기술표준이 제정될 가능성이 높음
- 기술표준은 고위험 AI에 대한 규제 요건 준수의 구체적 방법론으로 적용될 것으로 전망
 - EU 집행위원회 Joint Research Centre 2021년 보고서는 EU AI 규제 법안 요건과 기술표준의 대응 관계 제시

Requirements	Data and data governance	Risk management system	Technical data and Record keeping	Transparency and information to users	Human oversight	Accuracy, robustness, and cybersecurity	Quality management system
SDO							
ISO and ISO/IEC JTC1	ISO/IEC 25024; ISO/IEC 5259; ISO/IEC 24668;	ISO/IEC 4213; ISO/IEC 25059; ISO/IEC 24029-2	ISO/IEC 5338; ISO/IEC 5469; ISO/IEC 24368; ISO/IEC 24372; ISO/IEC 24668	ISO/IEC 24027; ISO/IEC 24028; ISO/IEC 5338; ISO/IEC 24368; ISO/IEC 24372; ISO/IEC 24668; ISO/IEC 4213		ISO/IEC 24027; ISO/IEC 24028; ISO/IEC 24029; ISO/IEC 5469	ISO/IEC 23894; ISO/IEC 38507; ISO/IEC 42001; ISO/IEC 25059
IEEE	ECPAIS Bias; IEEE P7002; IEEE P7003; IEEE P7004; IEEE P7005; IEEE P7006; IEEE P7009; IEEE P2801; IEEE P2807; IEEE P2863	IEEE P7009; IEEE P2807; IEEE P2846	ECPAIS Transparency; IEEE P7000; IEEE P7001; IEEE P7006; IEEE P2801; IEEE P2802; IEEE P2807; IEEE P2863; IEEE P3333.1.3	ECPAIS Bias; ECPAIS Transparency; ECPAIS Accountability; ; IEEE P7000; IEEE P7001; IEEE P7003; IEEE P7004; IEEE P7005;	ECPAIS Accountability; ECPAIS Transparency; ; IEEE P7000; IEEE P7006; IEEE P7014; IEEE P2863	ECPAIS Transparency; IEEE P7007; IEEE P7009; IEEE P7011; IEEE P7012; IEEE P2802; IEEE P2807; IEEE P2846; IEEE P2863;	IEEE 2801; IEEE P2863; IEEE P7000

(3) 제한적 위험 – 투명성 요구

- **투명성 적용 대상(제52조) – 챗봇 등**
 - 자연인과 상호작용하는 경우 → AI 시스템과 상호작용한다는 점을 통지
 - 감성인지 시스템 또는 생체인식을 통한 분류 시스템 → 해당 시스템 동작 고지
 - 콘텐츠 생성 또는 조작 (딥페이크) → 자동적으로 생성되었다고 표시
 - 예외 – 범죄 탐지, 예방, 조사 등 및 표현의 자유

(4) 최소한의 위험 – 규제 없음

- 나머지 대부분의 AI 시스템에 대해서는 규제가 없음



AI 규제에의 동향과 대응

Rule vs Standard

- **룰(rule)과 스탠다드(standard)**
 - 룰 – 허용 행위와 허용되지 않는 행위의 경계를 명확히 규정
 - 고속도로 최고 속도
 - 스탠다드 – 행위에 적용되는 기본적, 추상적 원칙을 규정
 - 안전운전 의무

Rule

Standard

국내 AI 규제의 동향

- **룰(rule)의 비용**
 - 최적의 규제 수준을 입법자가 사전에 정해야 함
 - **높은 오류 비용** – 과잉 규제 또는 과소 규제에 따른 사회적 비용
 - **높은 결정 비용** – 사전에 적절한 규제 수준을 위한 연구와 사회적 논쟁 비용
- **스탠다드(standard)의 비용**
 - 사후적으로 사법 판단에 의해 구체적 상황에 맞게 유연하게 해석 – 오류비용, 결정비용 낮음
 - **높은 규제준수 비용** – 수범자가 규제 내용을 파악하기 위한 비용을 지출
- **EU AI 법안에 대한 평가**
 - 고위험 AI에 대한 ‘스탠다드’ 형태의 규제
 - 자기 적합성 평가 과정을 통해 AI 공급자가 규제준수 비용 부담
- **국내 사업자들은 ‘룰’ 형태의 규제에 익숙함 → ‘스탠다드’ 형태의 규율 준수 역량을 높여야!**