

연구동향

July 2021 No.1

안전한 데이터 활용을 위한 데이터 위험관리

이경선 연구위원

정보통신정책연구원 디지털경제사회연구본부

안전한 데이터 활용을 위한 데이터 위험관리

이경선 연구위원

정보통신정책연구원 디지털경제사회연구본부, leeks@kisdi.re.kr

요약

- 본고에서는 2020년 KISDI에서 수행한 ‘안전한 데이터 활용을 위한 데이터 위험관리 보고서’의 주요 내용을 소개함
- 해당 보고서에서는 안전한 데이터 활용을 위한 데이터 위험관리를 위해 정보주체 관점에서 데이터 위험을 정량적으로 진단할 수 있는 데이터 위험진단 프레임워크를 제시하고 금융/의료/유통분야 데이터 위험을 진단하였음
- 분야별 데이터 위험수준은 국민, 산업 전문가, 법제도 전문가를 대상으로 온라인 설문조사를 수행하여 측정하였으며, 기회와 위험을 고려한 균형있는 정책설계를 위해 개인정보 활용이 가져올 혜택과 위험, 분야별 데이터의 수집/개방·공유/활용수준, 데이터 위험/위험관리 수준 등을 함께 조사하였음
- 이를 통해 1) 국민들의 데이터 위험인식 현황을 파악하고, 2) 국민-전문가, 산업전문가-법제도 전문가 사이의 데이터 위험 인식차이를 분석하였으며, 3) 분야별 데이터 위험-관리수준을 진단하고 고위험군/중점관리 영역의 데이터 위험유형을 파악하여 4) 향후 정부의 데이터 위험관리 정책방향을 제시하였음

01 개요

- 본 고에서는 KISDI에서 2020년에 수행한 「안전한 데이터 활용을 위한 데이터 위험관리 보고서」*의 주요 연구내용을 공유하고자 함

* 이경선 외(2020), 「안전한 데이터 활용을 위한 데이터 위험관리 보고서」, 정보통신정책연구원, 정책연구 20-14

- (연구목적) 안전한 데이터 활용을 위한 위험관리를 위해 다음과 같은 두 가지 목적을 가지고 수행
 - 데이터 위험의 속성을 이해하고 데이터 위험을 정량적으로 진단하기 위한 데이터 위험진단 프레임워크 제시
 - 데이터 위험진단 프레임워크를 활용하여 우리나라 주요 분야의 데이터 위험수준을 측정하고 안전한 데이터 활용을 위한 정부의 선제적 대응방안을 모색
- (연구범위) 정보주체(처리되는 정보에 의하여 알아볼 수 있는 사람으로서 그 정보의 주체가 되는 사람) 관점에서 개인정보의 활용이 정보주체에게 가져올 위험을 식별하고 금융, 의료, 유통 분야에 대해 데이터 위험 및 관리수준을 진단하여 대응방안을 모색

표 1 | 데이터 위험진단 및 분석 과정

사전조사	▶ 문헌조사, 이해관계자 인터뷰 등을 통해 데이터 개방 및 활용에 따른 데이터 위험의 속성을 이해
위험식별	▶ 정보주체의 관점에서 데이터 관련 위험을 포괄적으로 진단할 수 있는 데이터 위험진단 프레임워크 제시
위험측정	▶ 데이터 위험진단 프레임워크에서 식별한 위험유형의 데이터 위험수준을 국민, 전문가 대상 설문을 통해 측정하고 분석
위험진단	▶ 데이터 위험 및 관리수준에 기반하여 민간, 정부의 대응이 시급한 고위험군, 중점관리 영역을 진단하고 정책적 시사점을 도출

02 데이터 위험진단 프레임워크

- 데이터 위험의 속성을 불투명성, 불확실성, 불공정성으로 구분하고 미국 국가표준기술연구소(NIST: National Institute of Standards and Technology)의 프라이버시 프레임워크¹를 참조하여 정보주체 관점의 데이터 위험을 9가지로 유형화

¹ NIST(2020. 1. 16.), NIST PRIVACY FRAMEWORK: A TOOL FOR IMPROVING PRIVACY THROUGH ENTERPRISE RISK MANAGEMENT, VERSION 1.0

I 표 2 I 데이터 위험의 속성 및 9가지 위험유형

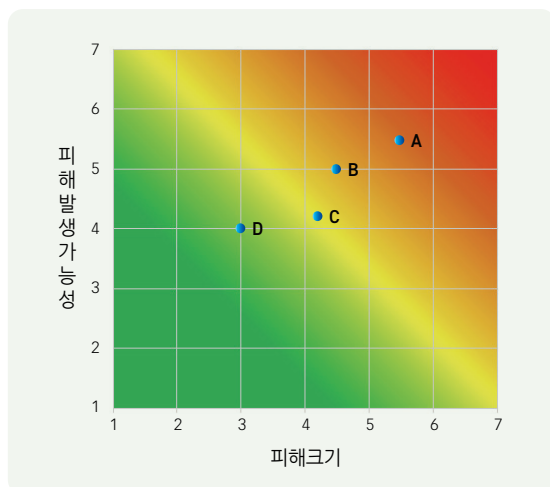
속성	데이터 위험		
데이터 수집/처리의 불투명성: 어떻게 적법하고 안전한 데이터의 수집, 처리, 활용을 담보하고 신뢰를 구축?	1	무단수집/이용	권한·예상을 넘어선 데이터 수집·이용의 위험
	2	부당한 접근제한	자기 정보에 대한 부당한 접근제한의 위험
	3	왜곡	부정확, 불완전한 개인정보의 생성, 이용, 유포 위험
데이터 처리의 불확실성: 어떻게 신속하게 위험을 정의하고 유연하게 대응?	4	보안취약	취약한 보안에 의해 데이터가 유출, 조작, 도용될 위험
	5	재식별	비식별처리 데이터의 재식별 위험
	6	예상치 못한 노출	맥락, 의미확장 등으로 원치않는 개인 민감정보, 사생활이 노출될 위험
데이터 활용의 불공정성: 어떻게 다양한 가치 간 균형을 맞추고 공정함을 유지?	7	과도한 정보수집/이용	이용목적 대비 과도하게 정보를 수집/이용할 위험
	8	감시	이용목적 대비 과도하게 정보추적, 모니터링할 위험
	9	편향된 추론	편향된 추론 데이터를 생성, 활용할 위험

- 데이터 위험관리를 위한 전략적 우선순위 파악을 위해 데이터 위험수준*을 진단하고 위험 수준에 따라 위험도를 고위험군, 위험군, 관심군, 안심군으로 구분

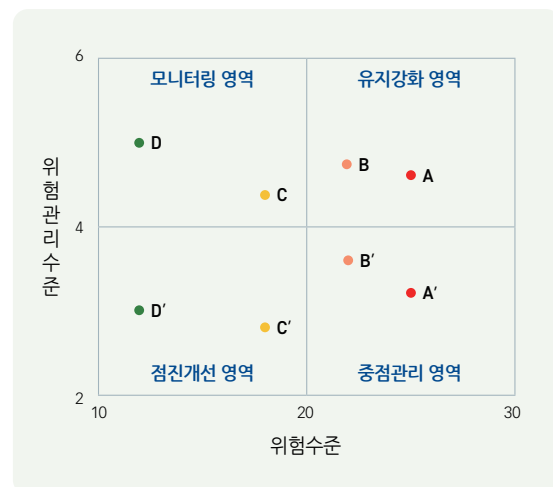
* 데이터 위험수준은 전통적인 리스크 관리 프레임워크에서 제시하고 있는 방식을 참조하여 ① 위험으로부터 실질적 피해가 발생할 가능성, ② 피해크기를 측정한다. ①과 ②를 곱해 계산

- 또한, 데이터 위험관리의 전략적 우선순위 결정을 위해 데이터 위험관리 수준을 측정하고, 위험 수준 대비 위험관리 수준을 비교하여 중점관리(위험 수준은 높지만 위험관리 수준은 낮은 영역), 점진개선(위험, 위험관리 수준이 둘 다 낮은 영역), 유지강화(위험, 위험관리 수준이 둘 다 높은 영역), 모니터링(위험 수준은 낮지만 위험관리 수준은 높은 영역) 정책영역을 파악

I 그림 1 I 데이터 위험도 구분



I 그림 2 I 데이터 위험관리 영역 구분



주: ● 고위험군, ● 위험군, ● 관심군, ● 안심군

피해발생가능성, 피해크기, 위험관리수준은 각각 7점 척도로 측정

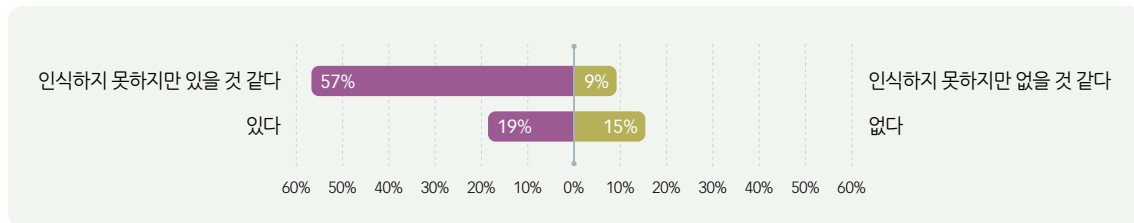
그래프의 A, A', B, B', C, C', D, D'는 각각 고위험군, 위험군, 관심군, 안심군에 속하는 예시임.

03 주요 결과

I 국민 데이터 위험 인식 결과

- 데이터 위험인식 조사 결과, 국민들은 데이터의 공정한 활용에 대한 회의적인 시선과 더불어 자신도 모르는 사이 데이터 위험에 노출되었을 가능성에 대한 높은 두려움을 드러냄
 - 개인정보 활용이 '취약계층의 서비스 접근성 개선'으로 연결될 것인가에 대한 동의수준은 42%
 - 데이터 위험의 경험여부에 대해 57%의 응답자가 "인식하지 못하지만 있을 것 같다"라고 응답

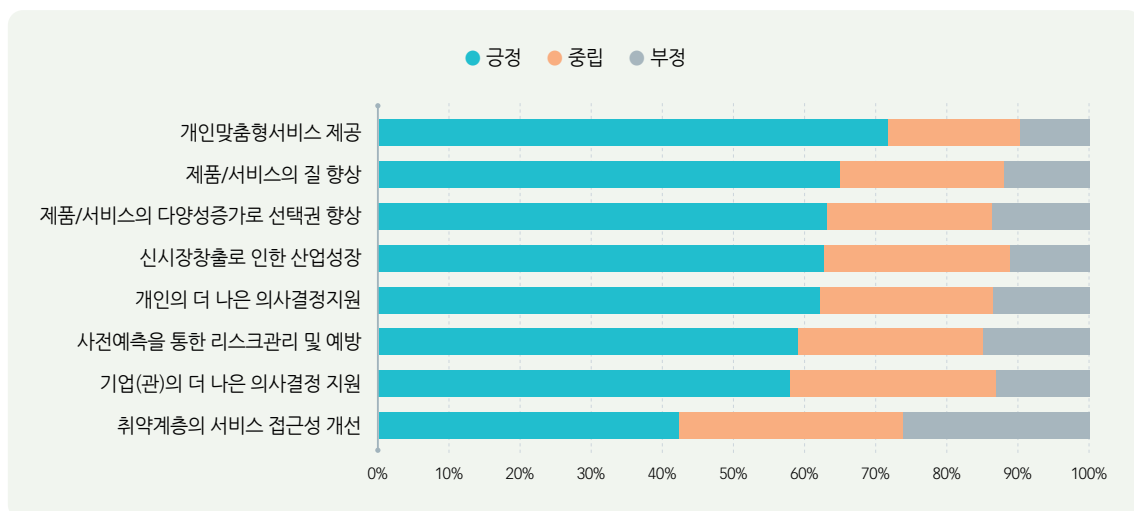
I 그림 3 I 데이터 위험의 경험여부



주: 9가지 위험유형별 데이터 위험 경험여부 응답의 평균값

- 그러나, 국민들은 개인정보 활용이 가져올 혜택에 공감하고 있었으며, 적절한 안전장치, 경제적 보상이 전제된 개인정보 제공에는 긍정적으로 반응
 - 개인정보의 활용이 가져올 혜택을 묻는 질문에서 '취약계층의 서비스 접근성 개선'을 제외한 모든 영역에서 50% 이상의 응답자가 긍정적이라고 응답

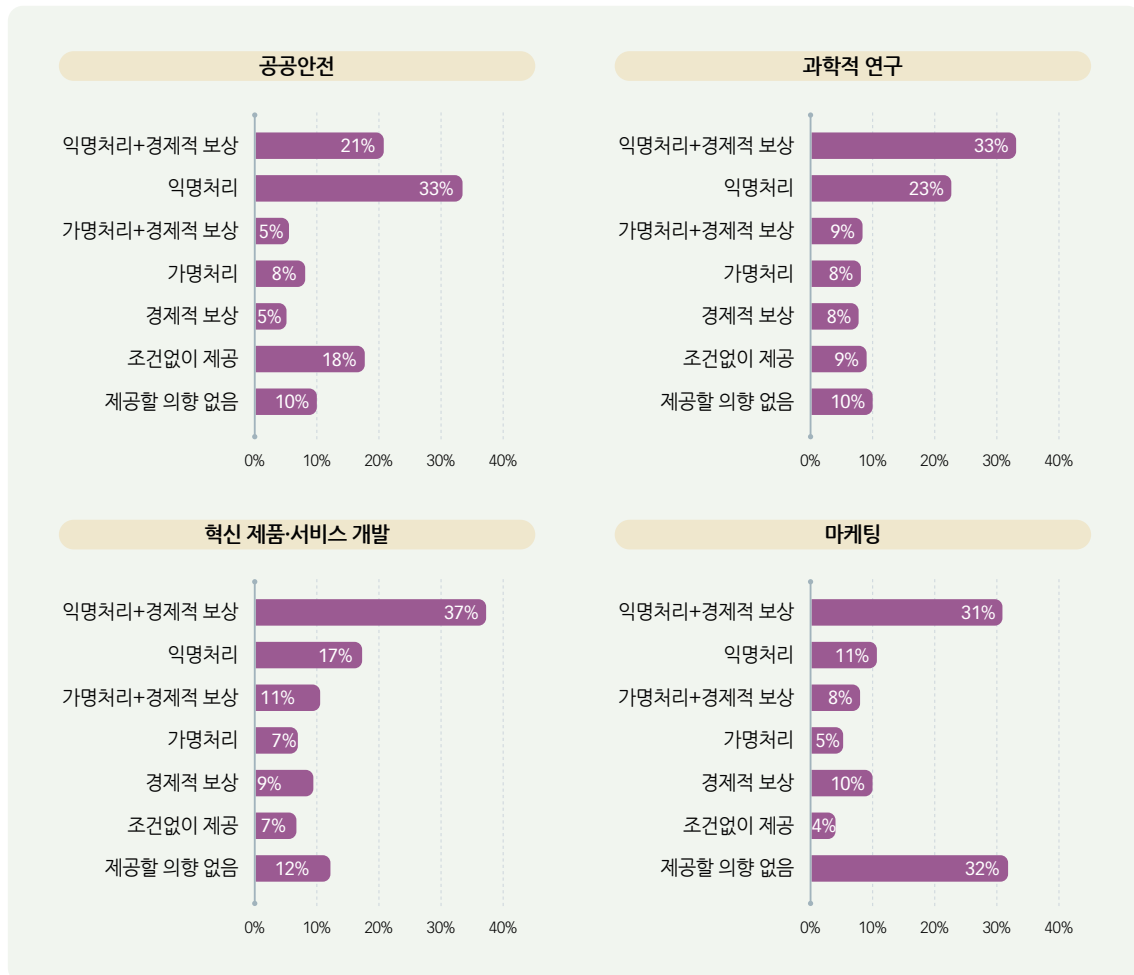
I 그림 4 I 개인정보 활용이 가져올 혜택에 대한 국민설문 응답결과



주: 0~7점 척도로 동의수준 측정 (긍정은 5점이상, 중립은 4점, 부정은 3점 이하를 집계)

- 또한, 개인정보 제공의향을 묻는 질문에 익명처리², 가명처리³, 경제적 보상이 각각 혹은 결합되는 방식으로 개인정보 제공을 요청받을 경우, “제공할 의향 없음”은 마케팅 목적의 경우 32%, 그 외 목적은 10% 수준으로 나타남

그림 5 | 목적별 개인정보 제공의향



주: 이 결과의 해석에는 두 가지 유의사항이 존재 1) 익명처리된 데이터는 더 이상 개인정보가 아니라는 점에서 기구축된 개인정보의 익명처리 후 사용은 개인의 동의없이 가능하며, 이에 “익명처리+경제적 보상”은 각각의 목적하에서 개인정보를 포함한 데이터를 처음 수집하는 시점에 사용가능한 옵션임 2) 개인정보 제공을 요청받았을 경우 제공할 의향이 있는지를 확인하기 위해 7가지 옵션 중 하나를 선택하도록 질문하여 얻은 결과로, 이 결과는 7가지 옵션 각각에 대해 제공의향을 물었을 경우에 비해 다소 보수적일 수 있음.

- 특히, “제공할 의향 없음”은 연령대가 낮아질수록 응답 비중이 낮아져 디지털 네이티브 세대*로 분류될 수 있는 20-30대에서 개인정보 제공에 가장 개방적인 것으로 나타남

* 1980년에서 2000년 사이에 태어나 디지털 기기, 디지털 기술과 함께 성장한 새로운 세대를 의미

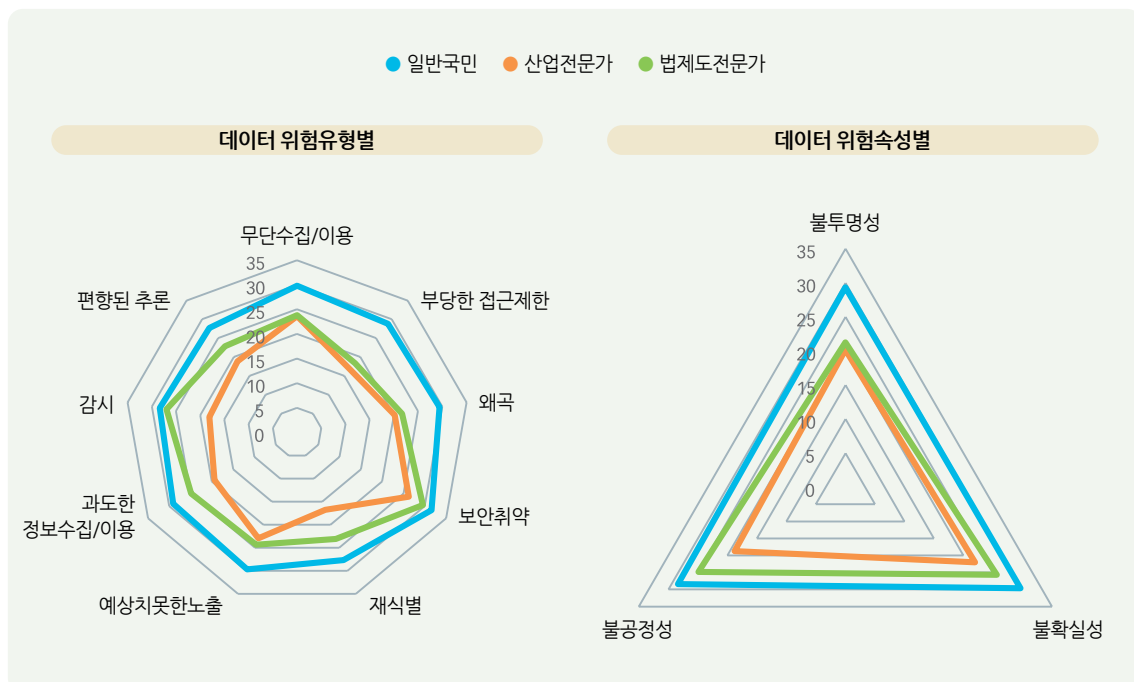
2 더 이상 특정 개인을 알아볼 수 없도록 처리하는 것을 의미하며 익명처리 후에는 시간, 비용, 기술 등 각종 수단을 동원해도 더 이상 개인을 알아볼 수 없음

3 개인정보의 일부를 삭제, 혹은 일부 또는 전부를 대체하는 방식 등을 통해 추가 정보 없이는 특정 개인을 알아볼 수 없도록 처리하는 것

I 데이터 위험인식 차이로 본 데이터 위험도

- 국민들이 느끼는 데이터 위험인식 수준은 전문가 대비 매우 높게 나타났으며, 특히 불투명성 관련 위험유형에서 차이가 크게 나타남
 - 개인정보 활용과 관련되어 국민들이 느끼는 두려움의 주요한 원인 중 하나가 정보의 비대칭성에 기인한다는 것을 의미하며,
 - 따라서 이를 완화하기 위해 데이터 수집·처리과정 전반의 투명성 제고, 개인정보자기결정권의 실질적 보장 등 정보의 비대칭성 해소를 위한 노력이 시급
- 전반적으로 외부 불확실성이 높아 정형화된 위험관리체계, 내부 관리로는 통제가 어려운 데이터 위험유형들의 위험수준이 가장 높게 나타남
 - 기술진화, 새로운 형태의 데이터 증가 등으로 데이터 생태계의 확장성, 불확실성은 향후 더욱 증가할 것이므로 변화하는 위험에도 작동가능한 위험관리체계의 구축이 중요
- 산업, 법제도 전문가 사이의 데이터 위험인식 차이는 불공정성과 관련된 데이터 위험유형에서 특히 높게 나타남
 - 산업 전문가는 현재 시점에서 현장에서 겪고 있는 실질적 위험을 중심으로, 법제도 전문가는 향후 대두될 잠재적 위험들을 중심으로 위험을 평가하였기 때문에 차이가 발생한 것으로 보이며,
 - 따라서 이러한 시각의 차이는 불공정성과 관련된 데이터 위험이 아직 산업현장에서 가시화되지는 않았으나 향후 다양한 산업에서 대두될 수 있어 선제적 대응이 필요함을 시사

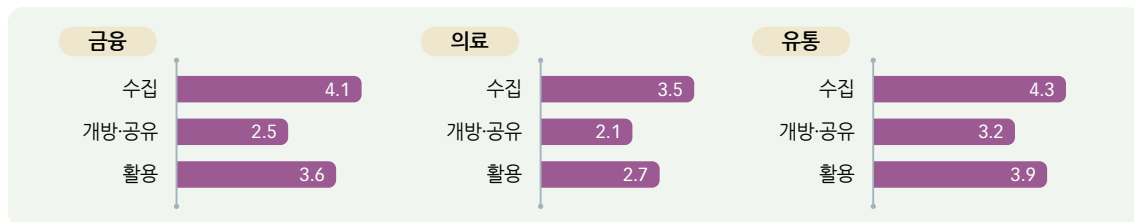
I 그림 6 I 그룹간 데이터 위험인식 차이



I 분야별 데이터 위험진단 결과

- 금융은 데이터 위험도가 높지만 대체로 위험관리가 적절하다고 평가됨, 단, 자동화된 의사결정 증가로 '편향된 추론' 위험에 대한 중점관리가 필요
- 의료는 데이터 활용 혜택에 대한 국민적 공감도가 가장 높으나 데이터의 수집·활용, 보호수준이 미흡하여 기회, 위험을 고려한 균형있는 대응 필요
- 유통은 데이터 수집·활용은 상대적으로 활발하나 데이터 위험인식이 부족한 중소·영세기업이 많아 위험인식 및 관리역량의 제고가 시급

I 그림 7 I 분야별 데이터 수집, 개방·공유, 활용수준



주: 1점 매우 낮음, 4점 보통, 7점 매우 높음

I 표 3 I 분야별 데이터 위험도 및 위험 대비 위험관리수준에 따른 전략적 우선순위

위험유형		금융		의료		유통	
		위험 ¹	관리 ²	위험	관리	위험	관리
무단수집/이용	권한·예상을 넘어서 데이터 수집·이용의 위험	●	유지강화	●	유지강화	●	유지강화
부당한 접근제한	자기 정보에 대한 부당한 접근제한의 위험	●	모니터링	●	모니터링	●	모니터링
왜곡	부정확, 불완전한 개인정보의 생성, 이용, 유포 위험	●	모니터링	●	점진개선	●	중점관리
보안취약	취약한 보안에 의해 데이터가 유출, 조작, 도용될 위험	●	유지강화	●	중점관리	●	중점관리
재식별	비식별처리 데이터의 재식별 위험	●	모니터링	●	점진개선	●	점진개선
예상치 못한 노출	맥락/의미확장 등으로 원치않는 민감정보, 사생활 노출 위험	●	유지강화	●	점진개선	●	중점관리
과도한 정보수집/이용	이용목적 대비 과도하게 정보를 수집/이용할 위험	●	모니터링	●	점진개선	●	모니터링
감시	이용목적 대비 과도하게 정보추적, 모니터링할 위험	●	모니터링	●	점진개선	●	유지강화
편향된 추론	편향된 추론 데이터를 생성, 활용할 위험	●	중점관리	●	점진개선	●	모니터링

주: 1 데이터 위험수준을 기준으로 ● 고위험군, ● 위험군, ● 관심군, ● 안심군

2 '중점관리'는 데이터 위험도가 위험군 이상 & 위험관리가 위험대응에 미흡한 수준을 의미

'점진개선'은 데이터 위험도가 위험군 미만 & 위험관리가 위험대응에 미흡한 수준을 의미

'유지강화'는 데이터 위험도가 위험군 이상 & 위험관리가 위험대응에 적절하거나 그 이상을 의미

'모니터링'은 데이터 위험도가 위험군 미만 & 위험관리가 위험대응에 적절하거나 그 이상을 의미

04 향후 대응방향

Ⅰ 데이터 위험의 속성별 위험관리 방안

- (투명성) 데이터 흐름 전반의 투명성 제고를 위해 정보주체의 실질적 권리보장, 기업의 책임성 강화, 사회전반의 신뢰자산 구축이 필요
- (신속·유연) 데이터 환경의 불확실성이 초래하는 위험에 보다 신속하고 유연하게 대처하기 위해 규모별, 수준별 차별화된 위험관리체계를 구축
- (공정성) 데이터가 우리 사회를 위해 책임있게 사용되도록 균형있는 데이터 정책 수립과 효과적인 국가 데이터 위험관리체계를 구축

Ⅰ 산업별·주체별 데이터 위험관리

- (금융) 데이터 편향성 평가 및 감사 등을 위한 연구 지원, 신용정보주체의 권리보호를 위한 지속적인 법제도 정비가 필요
- (의료) 동적동의 등 명시적 동의를 보다 쉽게 받을 수 있도록 제도를 개선하고 위험관리가 내재된 데이터 인프라 구축을 위한 지원책 마련
- (유통) 데이터 위험관리가 사업운영의 필수요소로 자리잡도록 분위기를 조성하고 개인정보 보호책임자 자격요건 전문화, 교육의무강화 등을 추진
- 건강한 데이터 생태계 구축을 위해 중장기적으로 정보주체를 포함, 생태계 내 주체들의 데이터 위험관리 역량의 제고가 필요

KISDI Perspectives 발간 내역



연번	제목	저자
2021-01-01	초점 망 중립성 가이드라인 개정의 배경과 주요 내용	라성현·정재윤
2021-01-02	정책동향 영국, 온라인 유해물 유통 회사에 대한 정책 방향 발표	전성호
2021-01-03	정책동향 영국의 유선전화 요금정책: BT의 자발적 요금 인상 제한 방안을 중심으로	윤도원
2021-02-01	초점 코로나19로 인한 소비패턴 변화	장재영
2021-02-02	초점 반도체산업 성장에서 플랫폼(Platform)으로서의 Foundry 중요성	김민식·이영종
2021-02-03	정책동향 코로나-19 확산에 따른 EU회원국의 대응동향	황해인
2021-03-01	초점 모바일 앱 카테고리별 시장구조 및 이용시간의 변화(2019년~2020년)	하승희·이채성
2021-03-02	연구동향 2020년 ICT기반 사회현안 해결방안 연구: 인공지능 사회정책 이슈와 대응 방향 고찰	조성은
2021-03-03	시장동향 터키의 방송시장과 한-터키 방송 콘텐츠 교류 현황	노은정
2021-04-01	정책동향 각국의 디지털서비스세 정책 동향	정연희
2021-04-02	초점 AI Multi-curation과 OTT 서비스 콘텐츠의 이용행태 변화	심홍진
2021-04-03	정책동향 금융분야의 마이데이터 제도 도입과 서비스 현황	이은민
2021-05-01	초점 EU 인공지능 규제안의 주요 내용과 시사점	이경선
2021-05-02	정책동향 디지털 환경에 대응한 독일 경쟁제한방지법 개정	이지현
2021-05-03	정책동향 PS-LTE 기반 재난안전통신망 구축 완료	김사혁
2021-06-01	초점 디지털 시대의 노년층: 포용 혹은 소외	전선민
2021-06-02	초점 미국의 플랫폼 규제 패키지 법안의 주요 내용 및 시사점	김현수·강인규
2021-06-03	초점 산업별 ICT 활용도 현황 분석	김정언·손녕선·신우철
2021-06-04	시장동향 Huawei에 대한 미국의 제재 영향 본격화에 따른 글로벌 스마트폰 시장 경쟁 변화 현황 및 전망	오정숙

KISDI PERSPECTIVES는 국내 외 정보통신방송 관련 주요 정책 및 시장 동향을 분석한 리포트입니다.

문의 : 노희윤 연구원 (정보통신정책연구원, ICT데이터사이언스연구본부, hyoooon@kisdi.re.kr, 043-531-4042)