

## 빅데이터의 위험 요소에 대한 고찰

천윤수\*, 박재경<sup>o</sup>

\*한국폴리텍대학 강서캠퍼스 사이버보안과,

<sup>o</sup>한국폴리텍대학 강서캠퍼스 사이버보안과

e-mail: cah3561@gmail.com\*, jakypark@gmail.com<sup>o</sup>

## A Study on Risks of Big Data

Yoonsoo Cheon\*, Jaekyung Park<sup>o</sup>

\*Dept. of Cyber Security, Korea-Politec University,

<sup>o</sup>Dept. of Cyber Security, Korea-Politec University

### ● 요약 ●

본 논문에서는 빅데이터의 활용이 확산되는 현대 사회에서 빅데이터의 수집, 관리, 이용 등에서 나타날 수 있는 문제를 확인하고 그 문제에 대한 기존의 대응 방법과 보완점을 시사한다. 빅데이터의 위험성은 개인 정보 유출, 디지털 디바이드, 편향성과 신뢰성, 의존성과 통제 가능성 등이 있다. 해당 문제는 빅데이터의 보편화가 가속될수록 큰 규모의 사회적 문제로 대두될 가능성이 높다. 이를 보완하기 위한 대응 방법을 크게 기술적 대응, 법적 대응, 사회적 대응으로 나누어 알아보고 각 부분의 취약점을 분석하여 개선의 방향을 제시한다.

**키워드:** 개인정보(Privacy), 디지털디바이드(Digital Divide), 편향성(Bias), 의존성(Dependency)

## I. Introduction

### 1. 필요성

대규모 데이터를 수집, 저장, 분석하여 새로운 부가가치를 창출할 수 있는 빅데이터의 활용은 사회 전반에 걸쳐 다양한 분야에 적용되어 있습니다. 빅데이터의 활용은 과거를 바탕으로 미래를 예측하고 그 미래를 준비하는 과정에서 편의성을 제공하고 효율성을 높이는 것에 기여합니다. 그러나 그 데이터는 개인 혹은 조직에 대한 고유한 정보를 담고 있기 때문에 수집 및 관리의 주체와 데이터 이용의 권한과 범위에 따라 문제가 발생할 여지가 있습니다.

### 2. 목적

수집 데이터와 수집 대상 그리고 활용이 포괄적인 빅데이터의 특성상 빅데이터 보안상의 문제가 발생하거나 법적인 체계에 맹점이 발생한다면 그 문제는 사회적인 문제의 규모로 커질 우려가 있습니다. 따라서 빅데이터의 위험성을 파악하고 이를 관리하는 방안이 필요합니다. 본 연구는 빅데이터의 안전하고 적절한 활용을 위한 방안을 제시합니다.

## II. Preliminaries

### Risks

#### 1. 개인정보 유출

개인정보의 무단 사용, 누출, 악용 등으로 개인의 권리와 자유를 침해할 우려가 있습니다.

<Graph 1>로 볼 수 있듯이 매년 개인정보와 관련한 유출 침해 사고가 증가하고 있음을 알 수 있습니다.

#### 2. 디지털 디바이드

빅데이터의 활용 과정에서 소외되는 지역이나 계층이 발생할 우려가 있습니다.

이는 디지털 격차를 심화시키고 분배와 평등성을 저해할 위험이 있습니다.

<Chart 1>은 성별과 연령 그리고 소득에 따른 디지털 정보화 수준을 보여주고 있습니다. 특히 소득 수준에 따라 디지털 정보화 격차는 다른 비교 기준에 비해 큰 차이를 보여주고 있습니다.

### 3. 편향성과 신뢰성

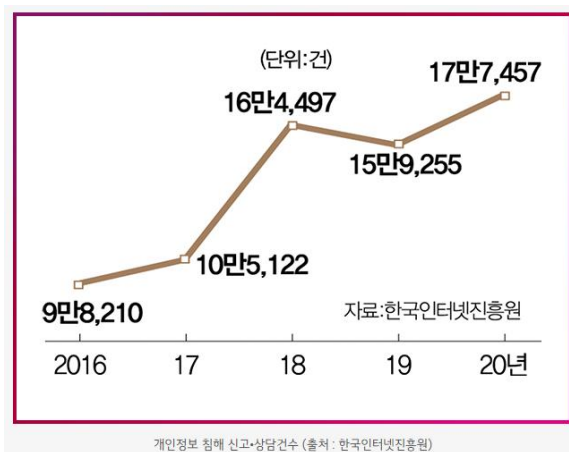
데이터 수집 표본의 선정, 특정 데이터를 제외한 분석 등의 방법으로 특정 목적을 위해 편향된 결과와 결정을 도출할 우려가 있습니다. 이는 빅데이터에 대한 신뢰성과 정확성에 치명적일 수 있습니다.

### 4. 의존성과 통제

빅데이터의 활용으로 얻을 수 있는 편의성은 개인이 해야 하는 사고와 결정의 영역을 편의와 효율의 명목 아래 도태시킬 우려가 있습니다.

또한 편의성으로 야기된 보편화는 특정 목적을 위해 특정 권한을 가진 개인이나 조직이 이용자의 의사결정에 간섭하거나 조작할 수 있는 우려가 있습니다.

<Graph 1>



개인정보 침해 신고·상당건수 (출처: 한국인터넷진흥원)

<Chart 1>

구분		종합	평균	역량	활용
성별	남성	83.8	97.2	75.7	82.5
	여성	77.6	94.7	66.5	76.5
연령	20대 이하	131.9	108.0	152.2	128.4
	30대	124.1	106.1	139.2	121.6
	40대	116.2	106.0	127.3	112.1
	50대	95.5	101.8	91.1	95.6
	60대 이상	71.1	92.9	57.0	69.9
월 가구 소득	100만원 미만	65.1	89.3	51.3	61.5
	100-199만원	66.7	91.7	50.1	65.5
	200-299만원	77.3	95.7	62.3	79.8
	300만원 이상	97.8	101.3	97.8	95.3
농어민 전체		78.9	95.7	70.6	78.8

출처: 2022 디지털 정보 격차 실태 조사, NIA, Ministry of Science and ICT

## III. The Proposed Scheme

### 1. 기술적 대응

빅데이터를 관리하는 기술적인 방법으로 익명화, 암호화 등이 있습니다. 그러나 암호화 알고리즘과 연산적인 부분의 취약성은 명백히 존재합니다. 또한 현재 데이터 보안에 관련된 인프라가 부족합니다. 따라서 강력한 암호화 알고리즘과 신뢰성 있는 인증 절차 메커니즘에 대한 개발과 연구가 필요하고 그에 맞는 지원과 투자를 활성화시켜야 합니다.

### 2. 법적 대응

개인정보에 대한 법적 규제는 개인정보 보호법, 정보통신망 이용촉진 및 정보보호 등에 관한 법률 등이 있습니다. 그러나 아직 많은 부분에서 법적 체계의 허점을 많이 드러내고 있습니다. 다양한 사례를 기반으로 개인 정보 유출과 악용이 파생시킬 수 있는 문제들에 대한 보완이 필요합니다.

빅데이터의 위험성에 대한 우리는 개인정보에 초점이 많이 맞춰진 경향이 있습니다. 데이터를 수집, 관리하고 활용하는 조직들이 신뢰성을 확보하고 편향성을 가지고 조작하지 않도록 하는 법적 제도가 부족합니다. 또한 조직들이 특정 목적을 가지고 차별적인 공급을 하여 발생할 수 있는 정보화 격차를 줄일 수 있는 법적 규제가 필요합니다.

### 3. 사회적 대응

<Graph 2>에서 볼 수 있듯이 대중들은 보통 개인정보의 유출이 피해자들에게 단기간에 직접적인 피해를 주지 않기 때문에 단일한 인식을 많이 가지고 있습니다. 반대의 경우도 마찬가지입니다. 인터넷을 기반으로 사회 전반에 걸쳐 다양한 첨단 기술들이 급격하게 변동하고 있는 시대에서는 교육도 그에 맞게 변화 해야 됩니다. 빅데이터에 대해 이해하고 위험과 문제를 바로 알 수 있는 교육을 기초로 점차 사회적인 인식을 개선해 나가야 합니다.

조직들은 자마다 그들의 목적을 가지고 빅데이터를 활용합니다. 그들의 이해관계는 복합적으로 서로 다른 조직들과 얽혀 있습니다. 따라서 문제가 발생한다면 그 문제는 여러 이해 관계자들과 이해관계가 얽혀 있습니다. 그렇다면 항상 책임의 문제가 대두됩니다. 때로는 이러한 책임의 문제가 문제 해결을 지연시키기도 합니다. 문제의 지연은 피해 상황을 악화시켜 가고 좋지 못한 선례를 남기기도 합니다. 조직들의 책임을 공유하고 협업을 통해 문제를 해결할 수 있는 가이드 라인을 만들고 따를 수 있도록 해야 합니다. <Graph 2>는 국내 조직들이 아직 개인정보에 관한 관리가 미흡하다는 것을 잘 보여줍니다.

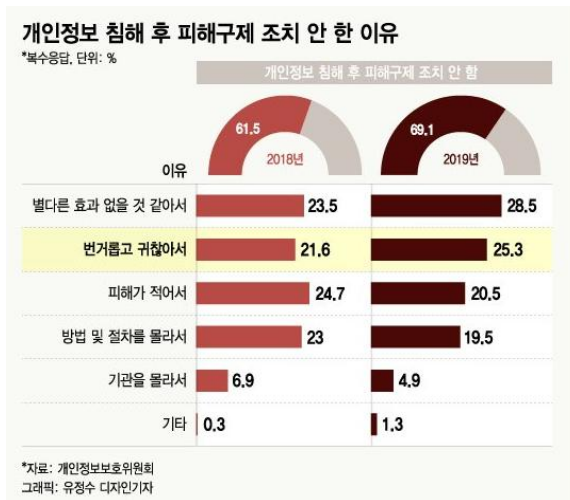
(Chart 2)

☐ 종합 결과 ※ 최고, 최저

구분(기관수)	평균(7797%)	중앙행정기관 (46개)	지방자치단체		중앙공공기관 (340개)	지방공기업 (1507개)	
			광역(17개)	기초(2267개)			
결과	84.3점	89.8점	82.8점		84.9점	83.8점	
			86.7점	82.5점			
비판 이유	양호	353개(45%)	28개(61%)	97개(53%)	79개(35%)	172개(51%)	657개(43%)
	보통	331개(43%)	16개(35%)	7개(41%)	117개(52%)	129개(38%)	627개(42%)
	미흡	95개(12%)	2개(4%)	1개(6%)	30개(13%)	39개(11%)	237개(15%)

출처: 2020년 공공기관 개인정보 관리수준진단 결과, 개인정보보호위원회

(Graph 2)



출처: Chanyoung Im, 국민 2명중 1명 개인정보 유출됐는데... "귀찮아서 신경안쓴다". 머니투데이. 2020.06.15.  
<https://news.mt.co.kr/mtview.php?no=2020061208391180982>

## IV. Conclusions

### 1. 빅데이터의 위험성과 대응 방안의 중요성

결론적으로 빅데이터는 우리의 삶과 사회 전반에 많은 혁신과 발전을 가져왔습니다. 그러나 빅데이터를 이용하는 과정에서는 여러 가지 위험성이 존재합니다. 이러한 위험성을 완전히 제거할 수는 없지만, 이를 최소화하고 관리하기 위한 방안이 필요합니다.

따라서, 빅데이터를 안전하게 이용하기 위해서는 이러한 위험성을 인식하고, 이를 관리하기 위한 방안을 지속적으로 모색하고 개선하는 것이 중요합니다. 빅데이터의 발전은 우리 사회의 발전과 끊임없이 연결되어 있기 때문에, 적극적인 대응과 노력이 필요합니다.

### 2. 한계점

빅데이터의 활용에는 여전히 한계점이 존재합니다. 빅데이터를 적극적으로 이용하기 위해서는 데이터의 품질과 정확성을 보장하고, 이를 활용할 수 있는 인력과 기술적인 인프라가 필요합니다. 또한, 빅데이터의 이용과 관리에 대한 법적 규제와 정책도 계속해서 개선되어야 합니다. 더불어, 개인정보 보호와 같은 윤리적 문제에 대한 해결책도 필요합니다.

이러한 한계점들은 빅데이터 활용의 발전을 방해할 수 있지만, 이러한 문제들에 대한 지속적인 연구와 대응책 마련을 통해 이를 극복할 수 있습니다. 따라서, 빅데이터 분석에 대한 연구와 기술 발전, 그리고 윤리적 문제와 법적 규제 등에 대한 논의와 개선이 계속되어야 할 것입니다.

## REFERENCES

- [1] Lane, J., V. Stodden, S. Bender, and H. Nissenbaum (2014). Privacy, Big Data, and the Public Good: Frameworks for Engagement. Cambridge University Press.
- [2] Milberg, S. J., S. J. Burke, H. J. Smith, and E. A. Kallman (1995). Values, personal information privacy, and regulatory approaches. Communications of the ACM 38 (12), 65-74.
- [3] Ohm, P. (2010). Broken promises of privacy: Responding to the surprising failure of anonymization. UCLA Law Review 57, 1701.
- [4] Anderson, S. and A. de Palma (2012). Competition for attention in the information (overload) age. RAND Journal of Economics 43 (1), 1-25.
- [5] UNESCO [66498] (2022). Minding the data: protecting learners' privacy and security. ISBN: 978-92-3-100525-1, Licence type: CC BY-SA 3.0 IGO [11819]
- [6] Ruth Gavison (Jan., 1980). Privacy and the Limits of Law. The Yale Law Journal. Vol. 89, No. 3, pp. 421-471 (51 pages), Published By: The Yale Law Journal Company, Inc.
- [7] Lee Dae Hee (2015). A Study on the Interpretation and the Scope of Personally Identifiable Information (PII). Korea Law Review, 2015, vol., no.79, pp. 165-207 (43 pages)
- [8] Ho, A. (2017). Advancing Educational Research and Student Privacy in the "Big Data" Era. Washington, DC: National Academy of Education