

개인정보 규제요인과 빅데이터 활용간의 관계에서 가명정보 결합의 매개효과 및 조절효과*

김상광**

요약

최근 빅데이터 활용의 영향요인으로 개인정보 규제요인과 가명정보 결합이 핵심 정책수단으로 등장하고 있다. 본 연구는 개인정보 규제요인과 빅데이터 활용의 관계에서 제3의 변수로서 가명정보 결합의 매개효과 및 조절효과를 실증분석하였다. 분석결과, 첫째, 개인정보 규제요인 중 개인정보 정의, 개인정보 동의, 법령위반 처벌강도 요인이, 그리고 가명정보 결합요인 중 결합 비식별성, 결합 가명정보 표준화, 결합 책임성이 빅데이터의 활용에 정(+)의 유의한 관계를 보였다. 둘째, 가명정보 결합 요인 중 결합 비식별성, 결합 가명정보 표준화, 결합 책임성이 개인정보 규제요인과 빅데이터 활용과의 관계에서 정(+)의 매개효과를 보였다. 셋째, 개인정보 규제요인과 빅데이터 활용과의 관계에서 가명정보 결합기관 유형인 자유형, 중개형, 지정형의 순서에 따라 조절효과가 다를 것이라는 가설은 기각되었다. 이상의 분석결과를 기반으로 개인정보 보호와 빅데이터 활용이 조화를 이루는 '착한규제'의 정책대안을 제시하였다.

주제어 : 개인정보, 규제요인, 빅데이터 활용, 가명정보 결합, 매개효과, 조절효과

The Mediating Effect and Moderating Effect of Pseudonymized Information Combination in the Relationship Between Regulation Factors of Personal Information and Big Data Utilization*

Kim, Sang-Gwang**

Abstract

Recently, increasing use of big data have caused regulation factors of personal information and combination of pseudonymized information to emerge as key policy measures. Therefore, this study empirically analyzed the mediating effect and moderating effect of pseudonymized information combination as the third variable in the relationship between regulation factors of personal information and big data utilization. The analysis showed the following results: First, among personal information regulation factors, definition regulation, consent regulation, supervisory authority regulation, and punishment intensity regulation showed a positive(+) relationship with the big data utilization, while among pseudonymized information combination factors, non-identification of combination, standardization of combined pseudonymized information, and responsibility of combination were also found to be in a positive relationship with the use of big data. Second, among the factors of pseudonymized information combination, non-identification of combination, standardization of combined pseudonymized information, and responsibility of combination showed a positive(+) mediating effect in relation to regulation factors of personal information and big data utilization. Third, in the relationship between personal information regulation factors and big data utilization, the moderating effect hypothesis that each combination institution type of pseudonymized information (free-type, intermediary-type, and designated-type) would play a different role as a moderator was rejected. Based on the results of the empirical research, policy alternatives of 'Good Regulation' were proposed, which would maintain balance between protection of personal information and big data utilization.

Keywords : personal information, regulation factor, big data utilization, pseudonymized information combination, mediating effect, moderating effect

Received Jun 30, 2020; Revised Jul 31, 2020; Accepted Aug 3, 2020

* This paper is edited from Dissertation of author. "A study on the Effect of Regulation Level in Personal Information on Big Data Utilization: Focusing on Data Combination"(2020).

** Research Fellow, Korea Social Responsibility Research Institute (tim2co@gmail.com)

I. 서론

4차산업혁명명은 단기적으로 인공지능의 등장, 장기적으로 과학기술의 유래없는 빠른 진전과 사회전반의 변혁을 의미한다. 4차산업혁명의 시대정신은 변동성(Volatility), 불확실성(Uncertainty), 복잡성(Complexity), 모호성(Ambiguity)을 특징으로 한다(4차산업혁명위원회, 2019). 4차산업혁명 시대에 인공지능(Artificial Intelligence)의 원료로 쓰이는 빅데이터는 새로운 부가가치가 무궁무진하게 창출되는 점에서 '21세기 원유'에 해당한다(Gartner, 2011).

최근 빅데이터 활성화의 영향요인으로 개인정보 규제요인과 가명정보 결합이 제약선과 핵심수단으로 등장하고 있다. 소비자 또는 국민의 개인정보(원본정보)를 그대로 수집해서 이용하고 유통하는 행위는 개인의 프라이버시를 침해하는 위법행위이다. 그러므로 누구의 정보인지 알 수 없도록 개인정보를 비식별조치 또는 가명조치를 취하고 이중 데이터를 연계하여 새로운 서비스와 부가가치를 창출하는 '가명정보 결합'이 빅데이터 활용의 핵심활동으로 부각되고 있다. 빅데이터 활성화는 공공정보·의료·통신·교육·복지서비스 등 영역이 다른 이중 데이터의 결합이 핵심 요건이므로 빅데이터 활용을 촉진하는 개인정보 규제요인과 가명정보·비식별정보 결합의 매커니즘을 검증하는 것이 선결요건이 된다.

세계 빅데이터 시장규모는 2019년 약 232조원 규모이며 매년 11.9%의 급격한 증가추세에 있다(IDC, 2019a). 한편 2019년 국내의 빅데이터 시장은 1조 6700억원 규모로서 세계시장의 0.72%에 불과하지만 연평균 16.3%씩 증가할 정도로 잠재력이 높다(IDC, 2019b).

한국은 빅데이터 활용의 꽃이라 불리는 '가명정보 결합(Pseudonymized Information Combination)' 또는 '데이터 결합'을 지원하기 위해 2016년 '개인정보 비식별조치 가이드라인'을 제정했다. 그러나 개인정보 비식별조치 가이드라인의 법적 성질이

행정규칙 또는 행정지도에 불과하여 비식별조치에 대한 법적근거가 취약하여 기업은 빅데이터 활용에 어려움을 겪고 있다. 2017년말에 시민단체는 법률의 규정이 아닌 동 가이드라인에 의해 비식별조치를 취한 기업 20곳을 검찰에 고발하여 한동안 기업의 비식별조치가 중단 된 사례가 있다. 정부는 데이터 경제 활성화를 위해 가명정보의 도입과 통계작성, 과학적 연구, 공익적 기록보존 등의 목적으로 정보주체의 동의 없는 가명정보의 활용 그리고 행정안전부, 방송통신위원회, 금융위원회로 분산된 감독기구를 개인정보보호위원회로 일원화하는 것을 내용으로 하는 개인정보보호법 개정안 및 정보통신망법, 신용정보법 등 데이터 3법을 국회에 제출하였다(2018.11). 그리고 산업계와 시민단체의 격렬한 논란 끝에 2020년 2월에 데이터 3법은 개정되었고, 2020년 8월 5일부터 시행되었다.

개인정보와 빅데이터에 관한 연구는 개인정보 보호 규제의 문제점과 혁신방안에 대한 법정정책적 연구(Kim, 2017; Cho, 2017; Park, 2016; Lee, 2019), 빅데이터 기술도입에 관한 영향요인에 관한 연구(Yang, 2015; Kang, 2017; Ko, 2019), 개인정보 비식별조치 등 가명정보 결합 및 안전조치에 관한 연구(Ko, et al., 2017; Cha, 2019; Lee, et al., 2017) 그리고 빅데이터 활용정책의 비교연구(Choi, et al., 2018; Park, 2016; Koo, 2017) 등 다양하다. 그러나 지금까지의 연구는 데이터 경제 활성화를 위해 엄격한 개인정보 규제를 혁신해야 한다는 법정정책적 규범연구가 대부분이었고, 관련 통계의 부족으로 개인정보 규제요인과 가명정보 결합 등 빅데이터 활용과의 인과관계를 분석할 실증연구는 거의 없었다. 그러므로 본 연구는 개인정보 규제요인과 가명정보 결합이 빅데이터 활용에 미치는 영향력과 제3의 변수로서 가명정보 결합의 매개효과 및 조절효과를 분석하고 정책적 대안을 제시한다.

II. 이론적 논의와 가설설정

1. 개인정보 규제요인과 빅데이터¹⁾ 활용의 관계

개인정보 규제요인의 개념과 정의에 대한 실증연구는 거의 없었다. Kim and Lee(2012)는 규제수준과 강도를 결정하는 기존의 선행연구를 정리하여 ‘규제강도 결정요인’을 제시하였다. Kim and Lee(2012)는 산업, 환경 등 규제영역별 공통의 규제요인을 발굴하여 <표 1>과 같이, 규제강도를 ‘행위완결의 불확실성과 준수부담 및 처벌강도의 종합적 수준’이라고 개념화하였다. 그러므로 ‘규제강도 결정요인’을 개인정보 규제요인에 적용하여 정의하면, 개인정보 규제요인은

‘개인정보에 관한 의무행위의 불확실성과 의무 준수부담 그리고 법령위반시 처벌강도의 종합적 수준’이라고 할 수 있다(Kim & Kim, 2020b).

이러한 범주에 근거하여 대표적인 개인정보 규제요인의 후보군으로 가명정보 등 활용제도 도입, 개인정보 비식별화 근거 명확성, 데이터 결합규정 명확성, 개인정보 정의 명확성, 개인정보 동의 용이성, 개인정보 감독기구 통합, 개인정보 법률간 정합성, 법령위반시 적정 처벌강도, EU GDPR 비교시 적정 과징금 순서로 조사되었다. 최종적으로 개인정보 정의 명확성, 개인정보 동의 용이성, 개인정보 감독기구 통합성, 법령위반 적정 처벌강도를 도출하였다(Kim & Kim, 2020a; 2020b).

〈표 1〉 규제강도 결정요인 연구
(Table 1) Research on Regulatory Strength Determinants

Researcher	Regulation Sector	Regulatory Strength Determinants	Regulatory Strength Criteria		
			Uncertainty	Compliance Burden	Punishment Intensity
Ogus(1991)	Regulation Classification	Degree of Gov't Intervention	○		
Jae-hong Kim(2002)	Regulatory Strength Classification	Prohibition, National Discretion, National Consent	○		
Doo-rae Kim(2007)	Industry Safety Regulation in U.S.	Degree of Violation, Possibility of being caught			○
Mi-hong Lee, · Jeong-sik Kong(2003)	Corporate Location Regulation in Environment Policy	Degree of Violation			○
		Cost of Compliance		○	
		Possibility of being caught			○
Seon-hwa Lee(2009)	Initial Allocation of Pollution Emission	Cost of Regulation Compliance		○	
Doo-rae Kim(2010)	Environment Regulation	Fine Scale			○
In-kwon Lee et al.(2010)	Corporate Cartel	Possibility of being caught, Level of Punishment			○

source: Kim&Lee(2012) research re-citation

1) 본 연구에서 개인정보와 빅데이터, 데이터는 구분하지 않고 동일한 의미로 사용한다. 그 이유는 IDC 보고서(2017)에 따르면 빅데이터는 소비자 및 생산자의 끝단에 있는 스마트폰, 카메라, 웨어러블기기, IoT센서에서 나온 Endpoint 단계의 데이터의 비중이 약 70%를 차지하여 가장 비중이 높으면서 대부분이 개인과 연결된 정보이기 때문이다(Cho, 2017; Kim & Kim, :2020a, 2020b).

구체적으로 살펴보면, 먼저 개인정보 정의 명확성은 개인정보와 비개인정보 구분의 불명확성(Cho, 2017), 데이터 결합의 근거 모호성(Ko, et al., 2017), 비식별 조치 가이드라인의 법적근거 부족(Lee, 2017), 가명정보의 활용제도 미도입(Park, 2016) 등이 있다. 개인정보 동의 용이성은 엄격한 정보주체의 개인정보 사전 동의(Kwon, 2015), 개인정보 수집·이용시 동의요건과 제3자 제공시 동의요건의 차이(Kim, 2017), 행위시 원칙금지 Positive 입법방식으로 사후동의 어려움(Lee, 2019) 등이 있다. 또한 개인정보 감독기구 통합은 분야별 감독기구 분산과 관할범위의 중복(Kim, 2017), 개인정보보호법, 정보통신망법, 신용정보법 등 개인정보 법률간 적합성 부족(Heo, 2014), 감독기구의 개인정보 활용기능 미도입(Kim, 2017) 등이 있다. 한편, 법령위반 적정 처벌강도는 법령위반시 형사별 위주의 처벌(Kim, 2019), EU의 개인정보보호법인 GDPR에 비해 높은 형사벌과 낮은 과징금(Kim, 2019), 민간협회 자율규제 미흡(Shin, 2018), 개인정보 유출시 엄격한 신고통지제도 운영(Lee, 2017), 클라우드 서비스 이용 시 개인정보 위탁·수탁 규제(Korean Chamber of Commerce, 2018)등을 세부 규제요인으로 들 수 있다.

다음으로 빅데이터 활용이론은 기술수용모델, 통합 기술수용모형, TOE(기술·조직·환경) 프레임워크가 있다. 먼저, 기술수용모델(TAM, Technology Acceptance Model)은 빅데이터 등 신기술 도입의 수용요인으로 '지각된 유용성'과 '사용 용이성'은 태도와 수용의도를 거쳐 최종적으로 수용행위에 영향을 미친다. Yang(2015)은 TAM을 이용하여 개인정보 비식별화의 기술성숙도가 빅데이터 활용행위에 미치는 영향을 실증분석하였다. 한편, Park(2016)는 TAM의 핵심변수인 지각된 유용성과 지각된 이용성을 통해 빅데이터 활용에 미치는 영향요인으로 법제도·행태·기술 등 데이터 인프라요인과 데이터 설계관리·DB관리·활용관리 등 데이터품질 요인을 실증분석하였다.

다음으로 Venkatesh, et al.(2003)는 통합기술수

용이론(UTAUT, Unified Theory of Acceptance and Use of Technology)을 발표하였다. 즉 Venkatesh, et al.(2003)는 TAM이 신기술 도입에 대한 조직차원의 지각과 인식에 치중하였다고 비판하고, 신기술에 대한 개인의 수용행위모델로서 UTAUT를 제시하였다.

한편 기술·조직·환경 프레임워크(TOE Framework)는 TAM과 UTAUT의 한계점을 극복하기 위하여 정보기술 활용요인으로 기술요인(Technology) 외에 조직요인(Organization)과 환경요인(Environment)을 도입하였다. 즉 TAM이 신기술 수용의 심리적 효과에 치중하여 실제 수용행동에 대한 설명이 부족하고, UTAUT는 개인의 신기술 수용에 초점을 맞추고 있어 조직차원의 기술수용을 포섭하기 어렵다고 비판한다. 그러므로 실제 조직차원의 행태변화 설명에 적합한 통합모형으로서 TOE Framework가 등장하였다(Kim & Kim, 2020a).

기술수용이론, TOE 프레임워크, UTAUT 등 빅데이터 수용에 대한 선행연구는 신기술 사용의도와 활용의지, 사용행동을 종속변수로 채택하고 있다. Venkatesh, et al.(2003)는 기술수용이론(TAM)에서 빅데이터 활용에 미치는 영향요인을 분석하였고, 종속변수로서 신기술 사용의도를 채택하였다 Chan, et al.(2013)는 빅데이터 사용행동을 빅데이터 활용에 대한 변수로 선정하였다. Yang(2015)은 신기술 이용활성화 의도를 종속변수로 선정하였고, Kang(2017)은 빅데이터 활용의 측정요인으로 활용의지와 활용정도를 채택하였다. 한편 Lee(2016)은 빅데이터 사용행동을 최종 종속변수로 채택하였고, Woo(2018)는 비식별화 기술의 사용행동을 신기술 도입과 활용변수로 선정하였다. 최근 Ko(2019)은 신기술 사용의도를 종속변수로 사용하였다. 이상의 논의를 종합하여 본 연구에서는 빅데이터 활용의 구성요소로서 '빅데이터 활용행동'을 선정한다.

한편, 개인정보 규제요인과 빅데이터 활용의 관계는 일반적으로 개인정보 정의규제, 동의규제 등 규제요인

이 명확하면(규제수준이 낮으면) 빅데이터 활용은 증가할 것이라는 규범적 논의가 다수 제기되었다. 즉 빅데이터 활용과 개인정보보호 규제요인은 상충·대립 관계(Trade-off)를 갖는다. 우선, 개인정보 보호 기본 원칙과 빅데이터 활용은 대립된 긴장관계를 갖는다. 개인정보의 최소수집원칙, 목적외 활용금지 원칙, 데이터 품질보장의 원칙, 투명성의 원칙은 빅데이터와 인공지능의 특성상 최대한의 대용량 데이터 수집과 위배된다(Yoon, 2018). 또한 빅데이터는 고정된 목적이 없이 대량으로 수집된 이후에 데이터의 가치와 로직, 데이터 분석의 패턴 등이 사후에 분명해지고 분석결과에 따라 가치가 달라진다는 점, 빅데이터 환경에서는 정보주체가 자신의 데이터가 자동 수집되어 수집된다는 사실 자체를 인지하기 어려운 점 등 데이터 수집단계부터 기술적 특성으로 인해 데이터 활용과 개인정보 보호는 상충하는 목적을 추구하게 된다(Kim & Kim, 2020b). 한편, Kim(2014)은 빅데이터와 프라이버시는 높은 기회와 위험을 파생하는 패러독스(Paradox)적 성격을 가지게 되므로 사용자의 신뢰확보와 절차적 정당성이 필요하다고 주장한다.

예외적으로 개인정보 보호와 빅데이터 활용의 관계는 상보·조화 관계(Complementary Relationship)가 가능하다는 주장이 있다. Park(2016)은 빅데이터 분석 기술과 프라이버시를 상호대척관계로 볼 게 아니라, 개인정보보호를 빅데이터 분석의 내재적 가치로 인식하는(Big data with Privacy) 통합관계로 볼 것을 제안한다. Andreas Weigend(2017)는 ‘포스트 프라이버시’에서 데이터를 공유해 얻을 수 있는 가치는 프라이버시 침해보다 크므로 정보주체는 기업의 데이터 활용을 허용하는 대신 개인정보의 투명성(데이터 접근, 데이터 점검)과 주체성(데이터 수정, 데이터 지움, 데이터 실험, 데이터 이전)을 확립하여 보호와 활용의 조화가 필요하다고 한다. 그러므로 개인정보 규제요인과 빅데

터 활용은 기본적으로 상충·대립하는 관계를 가지며, 예외적으로 상보관계가 성립할 수 있다.

〈가설 1〉 개인정보 규제요인(정의 명확성, 동의 이성, 감독기구 통합성, 법령위반 적정 처벌강도)은 빅데이터 활용에 정(+)의 영향을 미친다.

2. 가명정보 결합과 빅데이터 활용의 관계

가명정보 결합(Pseudonymized Information Combination)은 우선 두 개 이상의 개인정보(원본 정보)를 함께 가져와서, 공통 대체키 값인 추가정보(Additional Information)의 결합이 없이는 누구의 정보인지 알 수 없도록 가명조치를 한다. 이어서 서로의 데이터를 연계하여 데이터 간의 관계를 밝히는 것으로 ‘가명정보의 이종간 결합’을 의미한다(Lee, et al., 2017; Kim & Kim, 2020b). 가명정보 결합은 가명조치-적정성 평가-사후조치의 단계로 이루어지는데 이때 가명조치는 가명정보 결합의 기술적 조치에 해당한다.²⁾

가명정보 결합의 구성요인에 대한 연구로서 TOE 프레임워크에 근거한 다양한 선행연구는 <표 2>와 같다. 먼저 Yoon(2014)은 기업의 빅데이터 활용요인으로서 시스템 도입, 데이터 유형(기술요인), CEO·직원 인식, 교육·인재육성(조직요인), 그리고 투자환경, 인프라 환경(환경요인) 등 TOE 요인의 영향력을 실증 분석하였다. 한편, Lee(2016)는 빅데이터 도입의 영향요인으로서 상대적 이점과 복잡성, 호환성, 보안 등 기술요인과 조직규모, 경영층 지원 등 조직요인 그리고 경쟁자 압력, 규정지원의 환경요인의 영향력을 분석하였다. Kang(2017)은 기술요인, 법제도요인, 데이터 요인이 행태요인을 매개하여 빅데이터 활용에 미

2) 개인정보의 가명조치는 데이터 값 삭제(data reduction), 가명처리(pseudonymization), 총계처리(aggregation), 범주화(suppression), 데이터 마스킹(data masking) 등을 통해 개인정보의 일부 또는 전부를 삭제하거나 대체함으로써 추가정보(임시 대체키값)의 결합이 없이는 특정개인을 식별할 수 없도록 하는 조치를 말한다(Ministry of the Interior and Safety, 2020; Woo, 2018).

치는 영향을 실증 분석하였다. Yoon and Oh(2018)은 보안우려, 기술호환성, 전문성, 업무적합성, 지식공유, IT교육 및 데이터 기반의사결정의 TOE 변수가 빅데이터 활용에 미치는 영향을 분석하였다. 또한 Woo(2018)는 데이터 결합(비식별화)의 기술요인(보안성), 조직요인(경영층 지원, 비용), 환경요인(경쟁자 압력, 법·정책 지원)이 빅데이터 사용행동에 미치는 영향을 실증 분석하였다. 한편, Lee and Jang(2018)은 해운 항만기업의 빅데이터 활용의도에 미치는 TOE 변수로서 상대적 이점·복잡성·호환성(기술요인), 조직규모·최고경영층 지원(조직요인), 정부정책·경쟁자 압력(환경요인)의 영향력을 실증 분석하였다.

이러한 TOE 프레임워크를 참고하여 본 연구는 가

명정보 결합의 구성요인으로 가명정보의 결합 비식별성, 결합 가명정보 표준화, 결합 책임성을 선정하였다(Kim & Kim, 2020b). 구체적으로 개인정보 보호기술, 보안기술, 기술적 호환성, 비식별조치 개발 등은 기술요인을 대표하는 가명정보 결합의 비식별성으로 분류하였고, 결합기관 유형, 업무적합성, 조직규모, 결합 가명정보의 최신성·호환성·정확성 등은 조직요인을 대표하여 결합 가명정보 표준화로 선정하였다. 그리고 경쟁기관의 결합에 대한 압력, 결합 가명정보의 외부 반출 승인 및 재식별 금지 등은 가명정보 결합의 책임성으로 구분하였다(Kim & Kim, 2020b).

일반적으로 가명정보 결합과 빅데이터 활용의 관계는 수단과 목적의 관계 또는 정(+)의 상관관계를 갖는

〈표 2〉 TOE 프레임워크 연구

〈Table 2〉 Research on TOE Framework

Researcher	Variables		
	Technology Factor	Organization Factor	Environment Factor
Lian, et al.(2014)	Data Security, Complexity, Compatibility, Cost	Executive Support, Sufficient Resources	Government Policy, Perceived Industry Pressure
Oliviera, et al.(2014)	Security Technology, Cost, Readiness	Executive Support, Organization Scale	Pressure of Competitor, Legal Support
Hee-jeong Yoon(2014)	Introduction of System, Data Type	CEO·Employee Awareness, Education·Talent Training	Invest Environment, Infra Environment
Seon-woo Lee(2016)	Relative Advantage, Complexity, Compatibility, Security, Cost	Organization Scale, Executive Support	Pressure of Competitor, Regulation Assistance
Jeong-muk Kang(2017)	Technology Development, Standardization Networking	-	Legal System Factor, Data Factor
Soo-young Yoon, Jae-in Oh(2018)	Security Concerns, Data Factor, Technological Compatibility	Professionalism, Work Suitability, Knowledge Sharing	IT Education, Decision-making Based on Data
Soon-kyu Woo(2018)	Security(Non-identification)	Executive Support, Cost	Pressure of Competitor, Legal Policy Support
Jun-pil Lee, Myeong-hee Jang(2018)	Relative Advantage, Complexity, Compatibility	Organization Scale, Top Executive Support	Government Policy, Pressure of Competitor

source: Ko(2019) research re-citation

다(Kim & Kim, 2020a). 가명정보 결합방식은 나라마다 보호와 활용의 공통수단이다. 그러므로 세계 각국은 가명정보 결합을 위한 비식별 기술의 개발과 민간의 데이터 브로커 양성, 데이터 거래소 설립 등 가명정보 결합정책에 대한 경쟁이 치열하다. 가명정보 결합을 통해 기존 데이터가 보유한 관계성과 영향관계를 추가로 밝혀 이전에 몰랐던 새로운 사실을 확정함으로써 부가가치를 증진시키기 때문이다. 따라서, 개인정보가 포함된 데이터의 연계와 융합을 통해 개인정보처리자는 새로운 부가가치를 창출할 수 있다.

그러므로 선행연구에 의하면 가명정보 결합은 빅데이터 활용과의 관계에서 목적과 수단관계, 원인과 결과변수의 관계를 갖는 것으로 나타났다. Kim and Song(2014)은 '빅데이터 기술수용의 초기특성 연구'에서 조직의 혁신성향이라는 독립변인이 데이터 결합의 하위변수인 '결합혁신의 필요성'이라는 변수와 유사하게 실증분석에 활용되었다는 것을 알 수 있다. Yang(2015)은 Davis의 기술수용이론과 Laufer의 프라이버시 계산이론을 기본모형으로 채택하여 데이터 결합 즉 개인정보 비식별화 기술수준의 성숙도가 유용성, 사용 용이성, 업무기술 적합성, 개인정보 침해위험의 다중 매개변수를 거쳐 빅데이터 활용의도에 미치는 영향을 실증 분석하였다. 이러한 조사결과는 개인정보 비식별화 요인이 곧 데이터 결합의 비식별성을 대표하는 요인으로서 데이터 결합이 빅데이터 활용에 긍정적인 인과관계가 있음을 나타내고 있다. Kang(2017)은 서울시 공무원 510명을 대상으로 지자체 차원의 빅데이터 활성화에 영향을 미치는 영향요인을 분석하였다. 연구결과, 데이터 요인(정확성, 객관성, 신뢰성, 데이터 관리)이 데이터 결합과 유사한 변수로서 빅데이터 활용에 영향을 미치는 선행연구로 볼 수 있다. Park(2016)는 기술수용이론(TAM)과 혁신확산이론 및 기술·조직·환경접근(TOE Framework) 이론을 기반으로 빅데이터 활용에 대한 결정요인을 탐구하였고, 데이터 품질요인은 크게 데이터 결합의 하위요인으로서 결합 가명정보의 표준화와 유사한 변인이

라고 할 수 있으므로 가명정보 결합이 빅데이터 활용에 미치는 영향을 실증분석 하였다. Woo(2018)는 조직차원의 기술·조직·환경(TOE) 프레임워크와 개인차원의 통합기술수용이론(UTAUT)을 결합한 통합모형(Integrated Model)을 제안하고, 금융회사 등에서 빅데이터 기반의 개인정보 비식별조치 사용의도와 사용행동에 영향을 미치는 영향요인을 검증하였다(Kim & Kim, 2020a). 이러한 분석결과는 TOE 변수 중 보안성, 최고경영층 지원, 경쟁자 압력 등이 데이터 결합의 기술적·관리적 접근과 유사한 변인으로 볼 수 있다. 그러므로 가명정보 결합과 빅데이터 활용은 정(+)의 관계를 갖는 것으로 추정할 수 있다.

여기에서 개인정보 규제요인과 가명정보 결합의 빅데이터 활용과의 관계를 검토할 필요가 있다. 결론적으로 가명정보 결합은 개인정보 규제요인과는 다른 기본적인 특성을 보이므로 개인정보 규제요인과 별도의 독립된 요인으로 분석한다. 먼저, 가명정보 결합요인은 개인정보 규제수준이 빅데이터 활용에 미치는 영향력을 매개하는 효과를 갖는다. 즉 개인정보 규제라는 제약적 환경속에서 빅데이터 활용을 가속화 시키는 중개효과를 갖는 유력한 변수가 바로 개인정보 비식별조치와 적정성평가 등 데이터 결합의 기술적·관리적 요인이다. 예를 들어, 개인정보 정의 규제가 엄격하여 개인정보와 비개인정보의 구별이 모호한 상황 속에서 데이터 결합의 비식별 조치와 재식별에 대한 안전조치를 거쳐 빅데이터 활용은 증가하게 된다. 또한 개인정보 감독기구 규제가 엄격하여 분야별 규제법령과 규제기구가 중복 혼란이 되는 경우에 가명정보 결합의 비식별조치와 데이터 결합의 책임성은 수범자의 거래비용과 순응비용을 낮추게 되어 빅데이터 활용이 증가하게 된다. 또한, 개인정보 규제수준의 빅데이터 활용 억제효과에 대하여 가명정보 결합을 통한 처방책의 활성화와 경로, 규제수준을 낮추기 위해 제도적 처방책으로서 가명정보 결합요인이 기능을 수행한다. 즉 개인정보 규제수준을 낮추어 빅데이터 이용을 활성화시키는 진흥책으로서 가명정보 결합은 중요한 역할을 수행할 수 있다. 예를 들

어, 엄격한 개인정보 동의규제를 완화하기 위하여 개인정보의 식별성을 제거하거나 완화시키는 비식별조치 또는 가명조치의 결합을 통해 빅데이터 활성화를 추진할 수 있다. 개인정보(원본정보)간 결합이나 제3자 제공시 사전동의 규제를 완화하고 우회하여 과학적 연구, 통계조사, 공익적 기록보존 등의 목적을 달성하기 위해 정보주체의 동의 없이 안전한 개인정보의 이용이 가능하다. 이상과 같이, 가명정보 결합은 개인정보 규제요인과 구별되는 원인-결과 관계 및 목적-수단 등의 독자적 특성을 가지는 것으로 추정이 된다(Kim & Kim, 2020a). 그러므로 본 연구에서는 가명정보의 결합을 개인정보 규제요인과 빅데이터 활용의 관계에서 제3의 변수로서 조건부 효과를 살펴본다.

<가설 2> 가명정보 결합(결합 비식별성, 결합 가명정보 표준화, 결합 책임성)은 빅데이터 활용에 정(+의 영향을 미친다.

3. 가명정보 결합의 매개효과 및 조절효과

1) 개인정보 규제요인과 가명정보 결합의 관계

개인정보 규제요인과 가명정보 결합의 관계에 관한 선행연구는 거의 없으나 개인정보 규제요인의 명확성과 가명정보 결합의 수용성은 정(+의 인과관계를 갖는다고 추정할 수 있다. 먼저 개인정보 정의규정이 명확하고 비식별화의 적용요건이 확실하여 가명정보의 활용과 결합데이터의 범위가 명확해지면, 결합에 대한 위험부담과 거래비용이 감소하여 데이터 결합은 증가하게 된다(Park, 2014; Lee, et al., 2017). 개인정보 규제요인 중에서 정의규제가 가명정보 결합에 미치는 영향을 살펴보면, 개인정보 정의규정이 모호하고 비식별화, 익명화에 대한 용어규정이 혼란한 경우에 비식별조치 가이드라인의 법적근거에 논란이 있어 위험부담과 거래비용이 증가한다. 그러므로 거래비용의 증가로 인해 개인정보의 식별성을 제거하여 데이터의 부가가치를 높이는 가명정보 결합은 감소하게 된다(Park,

2017; Lee, et al., 2017). 한편 Ko, et al.(2017)는 개인정보 보호를 위해 비식별성을 높이는 익명화 기법이 주로 사용되고 있으나, 익명화 기법은 상대적으로 유용한 정보를 다수 손실하는 경향이 있어 적절한 수준의 비식별화가 필요하다고 주장한다.

둘째, 개인정보 동의규제와 가명정보 결합의 관계는 일반적으로 동의요건이 명확하여 동의규제가 낮으면 결합 가명정보 확보가 용이해지고 결합 가명정보의 가공과 분석, 품질관리에 소요되는 비용을 낮추어 가명정보 결합은 증가한다. 다만, 예외적으로 개인정보 동의규제와 데이터 결합의 관계는 엄격한 개인정보 사전동의 규제에 동의를 획득하기가 어려워지면 기업 등 개인정보처리자는 동의획득에 따르는 거래비용과 위험을 회피하고자 가명정보 활용요구로 비식별화의 필요성이 증가하고 가명정보 결합의 기술적 요인에 치중할 수 있다(Ko, et al., 2017). 또한 엄격한 사전 동의 획득 등 경직적 제도운영은 마이데이터, 데이터 결합 등 정보주체 동의에 기반 하지 않은 대안을 추구하는 경향이 있다(Cha, 2019). 결론적으로 엄격한 개인정보 동의규제는 결합 가명정보의 확보를 어렵게 하고, 결합데이터의 품질관리 비용이 증가하게 된다(Cho, 2017; Kwon, 2015; Kim, 2017).

셋째, 개인정보 감독기구의 통합은 감독기구의 분산 중복으로 인한 규제 및 감독권한의 혼란과 수범자의 혼란을 방지하여 순응비용을 낮추고 비식별조치 등 가명정보 결합은 증가한다(Heo, 2014). 개인정보 감독기구 중복혼선 등 감독기구 규제와 데이터 결합의 관계를 살펴보면, 개인정보 감독기구가 분산 중복되어 있으므로 개인정보 규제 및 감독권한이 중복되어 수범자의 혼란을 야기하고 순응비용이 증가하여 비식별조치 등 가명정보 결합에 대한 책임성이 저하된다(Heo, 2014). 즉 규제수준이 높을수록 비식별조치 등 가명정보 결합의 자율성과 책임성은 낮아지게 된다(Kim, 2019; Choi, et al., 2018).

넷째, 법령위반 시 적절한 처벌강도는 기관의 법률이행의무와 안전성 조치에 대한 준수부담을 낮추어 가

명정보 결합이 증가한다. 반대로 형사벌 등 과도한 처벌강도는 법령위반의 소지가 있는 가명정보 결합을 꺼리게 된다. 지난 2017년 11월에 시민단체가 ‘개인정보 비식별조치 가이드라인’에 근거하여 기업 간 데이터 결합을 수행한 전문기관과 기업을 검찰에 고발하여 지금까지 데이터 거래가 중단된 사례는 과도한 처벌강도의 부작용을 보여준다. 개인정보 처벌강도 규제가 높아 과도한 처벌이 내려지는 경우에, 의무준수부담이 크게 증가하는 기업 등 개인정보처리자는 형사벌 등을 우려하여 조금이라도 법령위반의 소지가 있는 이중 간 데이터 결합을 꺼리게 된다. 개인정보 처벌강도 규제가 높으면 과도한 처벌이 내려지는 경우에 법령에 대한 수범의무가 증가하여 가명정보의 결합은 위축되기 쉽다(Ko, et al., 2017; Ko, 2019).

2) 가명정보 결합의 매개효과: 결합 비식별성, 결합 가명정보 표준화, 결합 책임성

제3의 변수가 매개효과(Mediating Effect)를 가지려면 Baron and Kenny(1990)의 매개효과 분석모형에 따라 먼저, 개인정보 규제요인이 빅데이터 활용에 영향을 미쳐야 한다($X \rightarrow Y$). 다음으로 가명정보 결합과 빅데이터 활용의 관계에서 인과관계를 가져야 하고($M \rightarrow Y$), 마지막으로 개인정보 규제요인이 가명정보 결합과 유의미한 인과관계를 가져야 한다($X \rightarrow M$). 이 경우에 개인정보 규제수준이 빅데이터 활용에 미치는 직접적 인과관계가 있으면 가명정보 결합은 부분매개(Partial Mediation) 효과가 있다고 한다. 한편, 개인정보 규제요인이 빅데이터 활용에 미치는 직접적 인과관계 없이 매개변수를 거쳐 종속변수에 영향을 미치는 경우에 가명정보 결합은 완전매개(Full Mediation)효과가 있다고 한다. 선행연구 검토결과, 개인정보 규제요인과 가명정보 결합이 빅데이터 활용에 미치는 인과관계가 있는 것으로 나타났다(Kim & Kim, 2020a). 그리고 개인정보 규제수준이 가명정보 결합에 미치는 영향에 대하여 인과관계가 있는 것으로 추정 가능하므로 실증분석을 통하여 가명정보의 매개효과 및 조절

효과를 검증할 필요가 있다. 이 경우에 개인정보 규제요인과 가명정보 결합의 인과관계가 확인되는 경우에는 가명정보의 결합의 부분매개효과(Partial Mediation Effect)가 존재한다.

〈가설 3〉 가명정보 결합(비식별성, 결합 가명정보 표준화, 책임성)은 개인정보 규제요인과 빅데이터 활용의 관계에 대해 정(+의) 매개효과를 가질 것이다.

3) 가명정보 결합의 조절효과: 결합기관 유형

조절변수는 독립변수가 종속변수에 미치는 영향 가운데서 영향력의 크기와 방향(Strength and Direction of Effectiveness Power)을 바꾸는 제3의 변수이다. 본 연구에서는 가명정보 결합요인 중 결합기관 유형에 주목하여 영향력의 크기를 검증한다. 가명정보 결합기관 유형은 가명정보의 결합에 대한 국가의 개입 정도와 결합기관의 자율성을 기준으로 자유형(Free-type, 시설자격과 기술을 갖춘 민간기관이 결합을 수행하도록 국가가 허가하는 유형), 중개형(Intermediary-type, 신뢰할 수 있는 제3자가 데이터 결합을 중개하고 국가가認可하는 유형), 지정형(Designated-type, 국가가 독점적으로 조건을 갖춘 기관을 결합기관으로 지정하는 特許형)으로 나눈다. 국가개입의 정도가 높은 결합기관 지정형에 가까울수록, 개인정보 규제요인의 빅데이터 활용효과는 감소하게 될 것이다. 즉 결합기관 유형이 지정형인 경우, 가명정보 결합의 국가개입 정도가 높고, 기업의 자율성을 낮추므로, 개인정보 규제수준의 영향력이 증가하여 가명정보 결합비용을 높이고 위험을 증가시키므로 결국 빅데이터 활용은 감소한다. 반대로 결합기관 유형이 국가 개입 정도가 낮고, 기업 자율성이 높은 자유형에 가까울수록, 개인정보의 규제수준의 영향력은 감소하여 가명정보 결합비용을 낮추고, 궁극적으로 빅데이터 활용은 증가한다. 결론적으로 가명정보 결합기관 유형이 개인정보 규제요인의 명확성을 낮추는 지정형일수

록 빅데이터 활용은 감소하고, 개인정보 규제요인의 명확성을 높이는 자유형일수록 빅데이터 활용은 증가할 것이다. 그러므로 가명정보 결합기관 유형은 개인정보 규제요인과 빅데이터 활용과의 관계에서 영향력의 크기를 높고, 낮게 변화시키는 조절효과를 가질 것이라고 추정할 수 있다.

구체적으로 결합기관 유형은 자유형, 중개형, 지정형 순서대로 기업 등 개인정보처리자의 자율성이 낮고 국가개입이 높아 개인정보처리자에게 선호되지 않을 것이라는 추정이 가능하다(Lee, et al., 2017; Yoon, 2018). 즉 개인정보 규제수준은 데이터 결합의 국가개입정도가 높고, 기업의 자율성을 낮추므로, 규제강도는 증가하게 되어 가명정보 결합비용을 높이고 위험을 증가시키므로 결국 빅데이터 활용은 저조해 질 것이다. 그러므로 가명정보 결합기관 유형이 자유형에 가까울수록 결합기관의 데이터 활용범위와 데이터 축적이 가능해 지므로 향후 데이터 브로커(Data Broker)로 발전해 나갈 수 있다. 특히 민간의 가명정보 결합기관 허용은 초기 데이터 결합시장의 발전에 결정적인 역할을 수행할 것이다.

〈가설 4〉 가명정보 결합기관 유형 중 기관자율성이 높고 국가개입이 적은 자유형, 중개형, 지정형 순으로 조절효과가 다를 것이다.

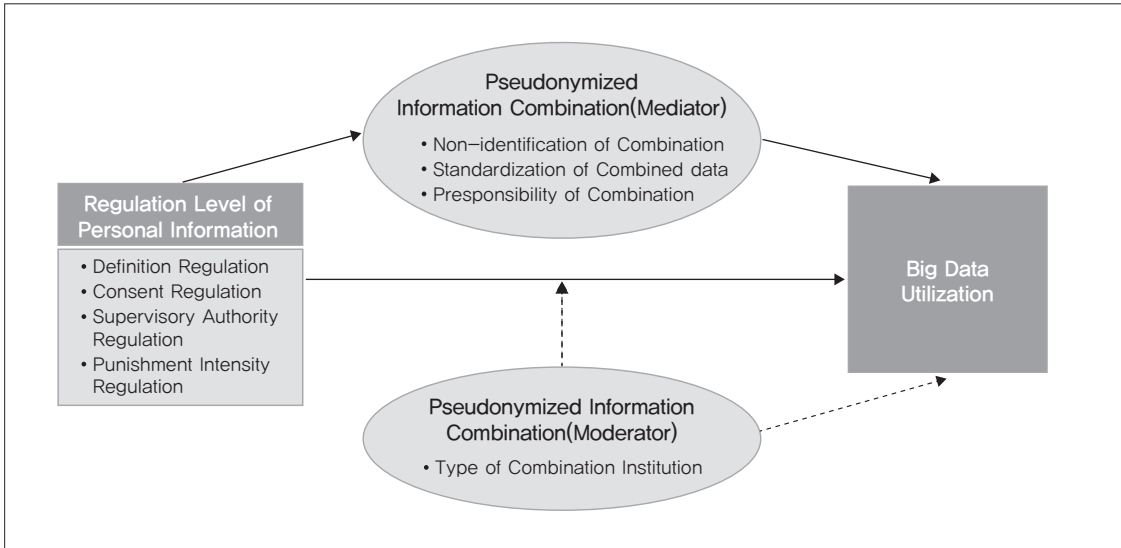
4. 선행연구의 비판적 검토

지금까지 개인정보보호와 빅데이터 활용에 관한 연구는 크게 네 가지 흐름을 가지고 진행되어 왔다. 첫째, 빅데이터 활성화를 위한 개인정보 보호제도의 문제점과 개선방안에 대한 법정정책적 규범연구이다. 둘째, 빅데이터 활용의 영향요인으로 공공데이터 개방, 보안 및 데이터 품질, 가명정보 결합, 기술수용이론(TAM), TOE 프레임워크 등 이론적 모형을 적용한 연구가 있다. 셋째, 데이터 결합연구로서 가명정보 이용, 비식별조치 기술개발, 방법론과 프로세스, 가명정보

결합에 대한 연구가 다수 있다. 그리고 넷째, 개인정보 보호 수준과 빅데이터 활용성과를 중심으로 국내 정책 현황을 고찰하고 해외 국가별 정책의 장단점을 논하는 비교연구가 있다.

기존의 연구는 다음과 같은 한계점을 보여 왔다. 첫째, 그간의 연구가 개인정보 정의 모호 및 비식별조치 가이드라인의 법적근거 미흡 등 개인정보보호제도의 문제점과 개선방안에 대한 법정정책적 연구에 치중하여 왔다는 점에서 규범적 연구에 머무르는 한계점을 보인다. 둘째, 기존연구는 TOE 프레임워크, 기술수용이론, 통합기술수용 모델(UTAUT) 등 빅데이터 시스템 도입과 활용에 미치는 영향요인을 중심으로 연구가 진행되었다. 최근에 빅데이터 활성화의 제약요인으로 개인정보 규제연구가 중심역할을 수행함에도 불구하고 개인정보 규제에 초점을 맞춘 빅데이터 영향요인 연구는 드물었다. 특히 가명정보 결합에 초점을 두고 심도 깊은 분석연구가 거의 없었다. 셋째, 부족한 개인정보 연구 중에서도 주로 개인정보 비식별조치 기술과 방법론, 비식별 조치 가이드라인의 법적 근거 등 가명정보 결합의 기술요인에 치중하여 연구가 진행되었다. 그러므로 가명정보 결합에 대한 관리적·절차적 연구는 그 중요성에도 불구하고 거의 언급되지 못하였다. 넷째, 그간의 연구는 개인정보 규제요인과 가명정보 결합에 대한 국내외 비교연구는 거의 이루어지지 않았다. 다섯째, 기존 연구는 법정정책적 규범적 연구위주로 이루어져서 개인정보 보호와 활용이 조화를 위해 적용가능한 정책적 대안이 크게 부족하였다.

그러므로 본 연구는 가명정보 결합이 개인정보 규제요인의 빅데이터 활성화 관계에 미치는 조건부 효과로서 매개효과 및 조절효과를 실증분석한다. 가명정보 결합의 구성요인으로서 결합 비식별성, 결합 책임성, 결합 가명정보 표준화, 결합기관 유형을 대상으로 기존의 ‘비식별조치’ 또는 ‘가명조치’라는 기술적 요인 중심의 접근방식과 더불어 데이터 결합의 관리적·절차적 측면에서 가명정보 결합의 매개효과와 조절효과를 살펴본다. 가명정보 결합에 대한 실증분석은 빅데이터



〈그림 1〉 연구모형
 〈Fig. 1〉 Research Model

활성화의 꽃이라고 불리우는 핵심적인 활동이면서 그동안 시도되지 않은 접근이라서 빅데이터 활용의 관점에서 새로운 시사점을 제공할 것이다. 일반적으로 빅데이터 활성화 요인을 탐구함에 있어서 데이터 개방지수나 전자정부 발전지수, 정보화 정도를 원인변수로 하여 연구해 왔으며, 개인정보에 대한 규제 강도와 규제정책을 독립변수로 한 연구는 매우 드물었고 실증 연구는 더욱 볼 수 없었다. 그러므로 본 연구는 개인정보 규제요인이 빅데이터 활성화를 촉진하는 유효한 정책수단이 될 수 있다는 단초를 제공한다. 동시에 전 세계적으로 규제수준이 높다고 평가를 받고 있는 우리나라 개인정보 보호의 규제수준을 되짚어 보고 유효한 정책대안을 제공할 것으로 기대된다.

Ⅲ. 연구설계

1. 연구모형

개인정보 규제요인은 빅데이터 활용에 영향을 미치

며, 제3의 변수로서 가명정보 결합이 매개효과 및 조절효과를 갖는다는 가설을 바탕으로 〈그림 1〉과 같이 연구모형을 설정하였다.

2. 변수의 조작적 정의

〈표 3〉과 같이, 독립변수인 개인정보 규제요인과 가명정보 결합 그리고 종속변수인 빅데이터 활용 변수의 개념을 측정 가능하도록 조작적으로 정의하였고, 변수에 따른 측정문항을 제시하였다. 자료 수집은 중앙부처, 지방자치단체 공무원과 공공기관 종사자 그리고 기업, 지원업체 등 민간의 관리자 등을 대상으로 이루어졌다. 설문조사는 행정안전부에서 실시하는 개인정보와 빅데이터 전문교육의 참석자를 대상으로 2019년 12월까지 진행하였다. 총 295부를 조사하였고 그 중 응답이 누락된 설문지 58부를 제외한 237부를 분석에 사용하였다. 설문조사의 통계분석은 SPSS 24.0과 AMOS 24.0을 사용하였다.

〈표 3〉 변수의 조작적 정의 및 측정 문항

〈Table 3〉 Operational Definition of Variables and Measurement Questions

Variables(Factors)		Operational Definition & Measurement Questions	Source
Regulation Factors of Personal Information	Certainty of Personal Information Definition	The degree to which personal information and pseudonym information is easily distinguished from non-personal information such as anonymous information	Jin-young Kim(2017), Seoung-Eun Cho(2017) etc.
		<ul style="list-style-type: none"> • The concept distinction between personal information and non-personal information is clear. • The definition of personal information is clearly defined and specified. • In Korea, the scope of personal information is clear, so it is easy to handle business. 	
	Easiness of Personal Information Consent	Degree of easiness of acquiring consent to personal information subject in legal requirements: collection and use of personal information, offering to third parties, provision of public affairs, urgent life and physical benefits, and purpose of statistics and academic research	Young-Jun Kwon(2015), Sang-moo Lee(2019) etc.
		<ul style="list-style-type: none"> • Upon the informed consent of the data subject, the personal information will be safely processed. • It is easy to obtain the consent of the data subject every time from the personal information controller's point of view. • It is easy to obtain the consent of the data subject every time from the information controller's point of view. 	
Integration of Personal Information Supervisory Authority	Degree of integration into one control tower; Personal Information Protection Committee(PIPC) dispersed by diverse Jurisdiction; Ministry of the Interior and Security(MOIS) takes on public institutions, the Korea Communications Commission(KCC) supervises online, and the Financial Services Commission(FSC) supervises financial agency	Personal Information Protection Committee(2018), Il-whan Kim(2017), Sooyeon Kim(2015) etc.	
	<ul style="list-style-type: none"> • The scope of regulatory jurisdiction for personal information supervisory authority by sector is clear. • It is easy to distinguish who the personal information supervisory authority is depending on the job. • Personal information supervisory authority's regulations have been clearly regulated, making it convenient. 		
Proper Criminal Punishment upon Violation of Laws	In case of violation of laws, proper criminal penalties such as imprisonment, fines are imposed instead of administrative penalties.	Yong-hak Kim(2019), Kyung-jin Choi(2018) etc.	
	<ul style="list-style-type: none"> • In case of violation of personal information protection laws, the intensity of punishment for imprisonment is appropriate. • In case of violation of personal information protection laws, administrative punishment, such as fines, is appropriate. • Korea has an appropriate level of punishment for violating laws. 		
Pseudonymized Information Combination	Non-identification of Combination	<p>Technical proficiency that eliminates the identification of personal information before heterogeneous data combining and the risk of re-identification of combined data and legal certainty regulated by law or statute</p> <ul style="list-style-type: none"> • In the case of combining pseudonymized information, non-identification measures lower the risk of re-identification. • In case of combining pseudonymized information, non-identification measures increase the effectiveness of data. • When combining pseudonymized information, the scope of use of the combined data should be clarified. 	Noh-hyung Park(2016), Hyeon-cheol Yang(2015), Soon-kyu Woo(2018) etc.

〈표 3〉 계속

Variables(Factors)		Operatioinal Definition & Measurement Questions	Source
Pseudonymized Information Combination	Standardization of Combined Pseudonymized Information	<p>Degree to which data types, formats, definitions, and components are processed into standardized data for easier combining of pseudonymized information</p> <ul style="list-style-type: none"> • The accuracy and clearness of pseudonymized information will be useful for combining data. • Interoperability of pseudonymized information will be useful for combining data. • Standardization of the quality of combined pseudonymized information contributes to the use of big data. 	Hyeon-sung Kim(2012), Jeong-muk Kang(2017) etc.
	Responsibility of Combination	<p>Strength of punishment for safety management in case of violation of pseudonymized information combination such as encryption of combined data; designation of security officer for combined data; and approval of export of combined data</p> <ul style="list-style-type: none"> • In the case of combining pseudonymized information, approval of external export of the combined data is required. • In the case of combining pseudonymized information, the combined data requires administrative, physical, and technical protection measures such as encryption. • In the case of combining pseudonymized information, safety measures are necessary to prevent personal information from being re-identified. 	Sung-wook Huh(2014), Jun-pil Lee·Myeong-hee Jang(2018) etc.
	Type of Combination Institutes	<p>Organizational type performing data combination activity such as non-identification and adequacy assessment when combining data and pseudonymized information (free-type, intermediary-type, designated-type)</p> <ul style="list-style-type: none"> • Defining the type of institutes combining pseudonymized information helps facilitate big data use with little government involvement and high-level of autonomy given to the institutes. • When combining pseudonymized information with other institutes, which institute type is most preferred? <ol style="list-style-type: none"> ① Free-type (the government permits public and private institutes based on facility qualifications, etc. to perform data combining) ② Intermediary-type (a trusted third party mediates data combining and is accredited by government) ③ Designated-type (patented by government exclusively as an agency for data combining) 	Eun-woo Lee, et al.(2017), Hak-soo Ko(2017) etc.
Big Data Utilization	<p>Intentions of use and actions such as business performance, continuous use, and intention to recommend using big data</p> <ul style="list-style-type: none"> • We are using big data and will continue to use it. • I am satisfied with my decision to use big data. • We will be happy to use big data. 	Venkatesh, et al.(2003), Soon-kyu Woo(2018) etc.	

source: Kim & Kim, s(2020b) re-citation

IV. 실증분석 결과

1. 표본특성

〈표 4〉와 같이, 응답자의 인구통계적 특성을 살펴보면 성별은 남자가 61.6%, 여자는 38.4%로 구성되었다. 가명정보 결합경험은 16.5%가 경험이 있다고 응답했고, 83.6%는 결합경험이 없었다. 직장은 중앙부처가 29.5%, 공공기관이 26.2%, 중소기업이 18.6%, 지방자치단체가 13.5% 등의 순으로 나타났다. 직종은 행정직이 46.4%, 기술직이 38.8%, 연구직 등이 3.0% 순으로 조사되었다. 직급은 7급 이하 또는 사원급이 57.4%, 6급 또는 차장급이 24.0%, 5급 이상 또는 팀장급 이상이 18.6% 이다.

2. 자료의 타당성 및 신뢰성 검증

일반적으로 연구모형의 타당성은 측정하고자 하는 개념을 얼마나 정확히 측정했느냐의 문제로서 요인 분석(Factor Analysis)이 주로 사용된다. 한편 신뢰성은 동일한 측정도구를 반복 측정하는 경우에 일관성 있는 측정결과를 얻는 것으로서 크론바흐 알파값(Chronbach's α)을 사용한다.

1) 탐색적 확인분석 및 신뢰도 검증

타당성 분석인 탐색적 요인분석(Exploratory Factor Analysis)은 정보의 손실을 막으면서 다수의 변수들을 동일한 요인으로 묶어 변수를 유형화시키는 방법이다. 동일한 요인을 추출하는 방법 중 가장 많이 이

〈표 4〉 응답자 표본특성
〈Table 4〉 Characteristics of Respondents

Type		Frequency (n=237)	%
Sex	Male	146	61.6
	Female	91	38.4
Combination Experience	Yes	39	16.4
	No	198	83.6
Office	NN	70	29.5
	Local Government	32	13.5
	Public Institute	62	26.2
	Big Company	25	10.5
	Small Company	44	18.6
	Self-employment	4	1.7
Occupation	Technical Job	92	38.8
	Administrative Job	110	46.4
	Research Job	35	14.8
Rank	Employee Level(7th)	136	57.4
	Team Manager Level(6th)	57	24.0
	Deputy Level(5th)	44	18.6

용되는 요인분석 모델은 주성분 분석과 공통요인분석이 있다. 본 연구에서는 관련된 변수를 축소, 압축하여 의미 있는 과정을 파악하기 위해서 주성분 분석을 사용하였다(Kim & Ahn, 2014). 그리고 요인을 회전하는 방법에는 변수의 독립성을 유지하는데 적합한 직각회전 방법(Orthogonal Rotation) 중 베리맥스 회전(Varimax Rotation) 방법을 사용하였다(Chae, 1997). 각 변수의 요인간의 상관관계의 정도를 나타

내는 요인적재량(Factor Loading)의 수용기준은 매우 높은 유의성을 갖는 $\pm .50$ 이상을 기준으로 선택하였으며(Kang, 1997), 각 요인이 전체 분산에 대해 설명할 수 있는 정도를 나타내 주는 고유값(Eigen Value)은 1 이상을 기준으로 하였다. 개인정보 규제요인, 빅데이터 활용, 가명정보 결합의 측정척도는 고유값이 1 이상인 8개의 독립된 요인으로 구분되고 있어 판별타당도(Discriminant Validity)가 확인되었다. 또한

〈표 5〉 탐색적 요인분석 및 신뢰도 분석결과
〈Table 5〉 Results of Exploratory Factor Analysis and Reliability Analysis

Construct	1	2	3	4	5	6	7	8	Communality	Cronbach's alpha
Definition Regulation	.109	-.100	.184	.307	.832	.140	.147	.083	.889	.891
	.077	-.041	.270	.338	.752	.105	.170	.205	.842	
Consent Regulation	.172	-.070	.205	.107	.099	.108	.886	.022	.775	.713
	.211	-.006	.022	.255	.301	.047	.642	.399	.895	
Supervisory Authority Regulation	.085	-.090	.186	.876	.210	.028	.142	.213	.912	.918
	.120	-.143	.082	.840	.339	.056	.133	.169	.927	
Punishment Intensity Regulation	-.031	.970	-.119	-.087	-.059	-.069	-.047	-.055	.975	.982
	-.036	.968	-.120	-.098	-.063	-.059	-.034	-.070	.977	
Big data Utilization	.128	-.084	.152	.278	.179	.065	.110	.857	.859	.715
	.117	-.292	.453	.299	.103	.057	.330	.467	.735	
Non-identification of Combination	.204	-.138	.802	.244	.088	.212	.190	.079	.866	.861
	.265	-.091	.708	.018	.427	.121	.060	.206	.912	
	.347	-.248	.584	.088	.343	.349	.083	.042	.834	
Standardization of Combined Pseudonymized Information	.832	-.009	.133	.030	.108	.214	.108	.202	.821	.913
	.876	.008	.253	.098	-.020	.153	.149	.017	.887	
	.888	-.093	.110	.109	.151	.216	.103	.021	.902	
Responsibility of Combination	.299	-.122	.094	.063	.210	.874	.014	.009	.925	.877
	.248	-.021	.255	.020	.015	.863	.138	.085	.899	
Eigen Value	7.461	2.329	1.798	1.126	.912	.828	.695	.582		
KMO(Kaiser-Meyer-Olkin)									.822	
Bartlett's Test						Chi-Square			3406.165	
						df(p)			153(.000)	

변수별 요인 적재량이 0.50 이상으로 각 요인에 비교적 높게 적재되어 있어 집중타당도(Convergent Validity)가 있다고 할 수 있다.

본 연구의 측정항목에 대한 요인분석 결과 <표 5>와 같이 8개의 요인이 추출되었다. 본 연구에서 개인정보 규제요인, 빅데이터 활용, 가명정보 결합 측정척도의 요인분석 결과 각각 요인1을 ‘개인정보 정의규제’로 명명하였으며 요인2는 ‘개인정보 동의규제’, 요인3은 ‘개인정보 감독기구 규제’, 요인4는 ‘법령위반 처벌규제’, 요인5는 ‘빅데이터 활용’, 요인6은 ‘가명정보 결합 비식별성’, 요인7은 ‘결합 가명정보 표준화’, 요인8은 ‘가

명정보 결합 책임성’으로 각각 구분하였다. 동일한 요인을 구성하는 표본의 수와 변수의 수가 적정한지 검증하기 위하여 KMO 검증과 Bartlett 검증을 실시한 결과, KMO(.822) > .5, Bartlett의 p값(.001)은 < .05를 만족하므로 탐색적 요인분석은 적정하다.

자료의 신뢰성은 일반적으로 크론바흐 알파(Alpha) 값이 0.6 이상이면 신뢰도가 높다고 보고 있다. 따라서 본 연구에서 사용한 측정도구의 신뢰성은 대부분 0.6 이상이기 때문에 척도의 신뢰성이 있는 것으로 확인되었다.

<표 6> 확인적 요인분석 결과

<Table 6> Confirmatory Factor Analysis Results

Construct	Non-standardized Estimates	S.E.	C.R	Standardized Estimates	Errors	AVE	C.R.(Construct Reliability)
Definition Regulation	1	0.059	17.079***	0.874	0.704	0.589	0.741
	1.015			0.921	0.419		
Consent Regulation	1	-.091	8.706***	0.851	0.864	0.469	0.634
	0.793			0.651	0.437		
Supervisory Authority Regulation	1	0.048	19.075***	0.993	0.392	0.685	0.812
	0.916			0.911	0.455		
Punishment Intensity Regulation	1	0.039	25.203***	0.993	0.031	0.920	0.959
	0.977			0.971	0.136		
Big data Utilization	1	0.099	10.410	0.689	0.406	0.550	0.708
	1.027			0.816	0.529		
Non-identification of Combination	1	-.065	14.957***	0.870	0.456	0.514	0.760
	0.974			0.812	0.698		
	0.928			0.777	0.759		
Standardization of Combined Pseudonymized Information	1	0.066	17.207***	0.825	0.466	0.633	0.838
	1.144			0.905	0.341		
	1.163			0.920	0.554		
Responsibility of Combination	1	0.068	13.530**	0.916	0.804	0.566	0.723
	0.915			0.853	0.397		

〈표 7〉 판별타당성 분석결과
 (Table 7) Discriminant Validity Results

Construct	Co-Relation							
	1	2	3	4	5	6	7	8
Definition Regulation	0.802							
Consent Regulation	0.251	0.809						
Supervisory Authority Regulation	0.581	0.431	0.784					
Punishment Intensity Regulation	-0.359	-0.375	0.388	0.779				
Big data Utilization	0.152	0.441	0.356	-0.484	0.799			
Non-identification of Combination	0.367	0.249	0.277	-0.337	0.375	0.759		
Standardization of Combined Pseudonymized Information	0.111	0.103	0.026	-0.097	0.077	0.136	0.828	
Responsibility of Combination	0.432	0.208	0.285	-0.238	0.156	0.329	0.134	0.727

2) 확인적 요인분석

탐색적 요인분석의 결과를 토대로 Amos 24 통계 패키지를 활용하여 확인적 요인분석(Confirmatory Factor Analysis)을 실시하였다. 〈표 6〉과 같이, 각 변수의 평균분산추출값(AVE)이 0.5이상, 개념신뢰도(C.R.)가 0.7이상으로서 총 8개 변수가 동질적 변수를 나타내는 집중타당성을 만족하였다(다만, 개인정보 등의 규제는 0.5에 근접하여 동일 개념으로 간주). 집중타당성은 동일한 구성개념을 나타내는 관측변수의 상관관계가 높아야 한다.

3) 판별타당성 분석

판별타당성은 서로 다른 잠재변수간 존재하는 측정치의 차이를 나타내며, 집중타당성과 달리 변수간 상관관계가 낮아야 한다. 판별타당성을 검증하기 위하여 변수에 대한 평균분산추출값의 제곱근 값이 개념 변수간의 상관계수보다 높은지 알아보았다. 즉 〈표 7〉과 같이, 독립변수, 매개변수, 종속변수를 형성하는 각 관측변수가 이질적이고 상관계수가 낮아 판별타당성이 있는 것으로 조사되었다.

3. 구조방정식 적합도 및 매개효과 분석

자료의 적합도를 검증하기 위하여 다양한 지수를 사용하였다. 그 결과를 보면 〈표 8〉에서 나타난 바와 같이, χ^2 값이 89.731로 나타났다. χ^2 값은 모델을 채택하는 경우, 충분조건이 아닌 필요조건이므로 다른 적합도 지수를 참조하여 적합도를 판정하였다. GFI는 .923, RMR는 .120, RMSEA는 .108, CFI는 .894, TLI는 .878, AGFI는 .856으로 수용 가능한 값으로서 적합도 기준 값에 가까운 결과가 나와 구조방정식 모형은 적합한 모형이라고 할 수 있다.

경로분석을 통해 독립변수, 매개변수, 종속변수 간의 인과관계를 직접효과, 간접효과, 총효과로 나누어 분석한 결과는 〈표 9〉와 같다. 경로분석은 공분산 또는 상관계수로 변수 간의 인과관계를 규명하는 방법이다. 경로 분석은 다중회귀분석이나 요인분석보다 복잡한 분석을 요구할 때 구조방정식 모형을 쉽게 분석해주는 방법으로써 다중회귀 분석에서 파악하기 어려운 직접효과와 간접효과, 총효과를 쉽게 파악하는 장점이 있다.

〈표 8〉 구조방정식 모델의 적합도 검정
 〈Table 8〉 Goodness-of-fit Test of Structural Equation Model

Goodness-of-fit Index		Optimal Model	Test Value
Absolute Fit Index	χ^2	>.05	89.731
	GFI	near 1.0	.923
	RMR	near 0	.120
	RMSEA	near 0	.108
Incremental Fit Index	CFI	near 1.0	.894
	TLI	near 1.0	.878
Parsimony Fit Index	AGFI	near 1.0	.856

〈표 9〉 경로분석
 〈Table 9〉 Path Analysis

Hypothesis Path		β	S.E.	C.R.	.P	B
Big data Utilization	← Regulation Level of Personal Information	.699	.066	10.575	.001	.567
Big data Utilization	← Definition Regulation	.403	.041	9.895	.001	.541
Big data Utilization	← Consent Regulation	.417	.041	10.173	.001	.552
Big data Utilization	← Supervisory Authority Regulation	.441	.039	11.438	.001	.597
Big data Utilization	← Punishment Intensity Regulation	-.262	.048	-5.443	.001	-.334
Big data Utilization	← Pseudonymized Information Combination	.494	.061	8.089	.001	.466
Big data Utilization	← Non-identification of Combination	.553	.052	10.237	.001	.555
Big data Utilization	← Standardization of Combined Pseudonymized Information	.291	.050	5.839	.001	.355
Big data Utilization	← Responsibility of Combination	.238	.055	4.907	.001	.270
Pseudonymized Information	← Regulation level of Personal Information	.510	.068	7.513	.001	.439
Pseudonymized Information	← Definition Regulation	.583	.065	8.343	.167	.556
Pseudonymized Information	← Consent Regulation	.463	.034	10.231	.087	.523
Pseudonymized Information	← Supervisory Authority Regulation	.442	.043	5.692	.001	.564
Pseudonymized Information	← Punishment Intensity Regulation	-.342	.056	-4.746	.213	-.412

개인정보 규제요인은 빅데이터 활용에 총효과 (.567), 직접효과(.449), 간접효과(.118)를 동시에 주고 있는 것으로 나타나 직접효과는 물론 매개변수를 통한

간접효과도 주는 것으로 나타났다. 매개변수의 간접효과의 유의성 정도는 부트스트래핑(Bootstrapping) 검증결과 p값이 .001로 유의하게 나타났다. 따라서 가명

정보 결합은 매개변수로 개인정보 규제요인과 빅데이터 활용의 매개변수로 영향을 미치는 것을 알 수 있다.

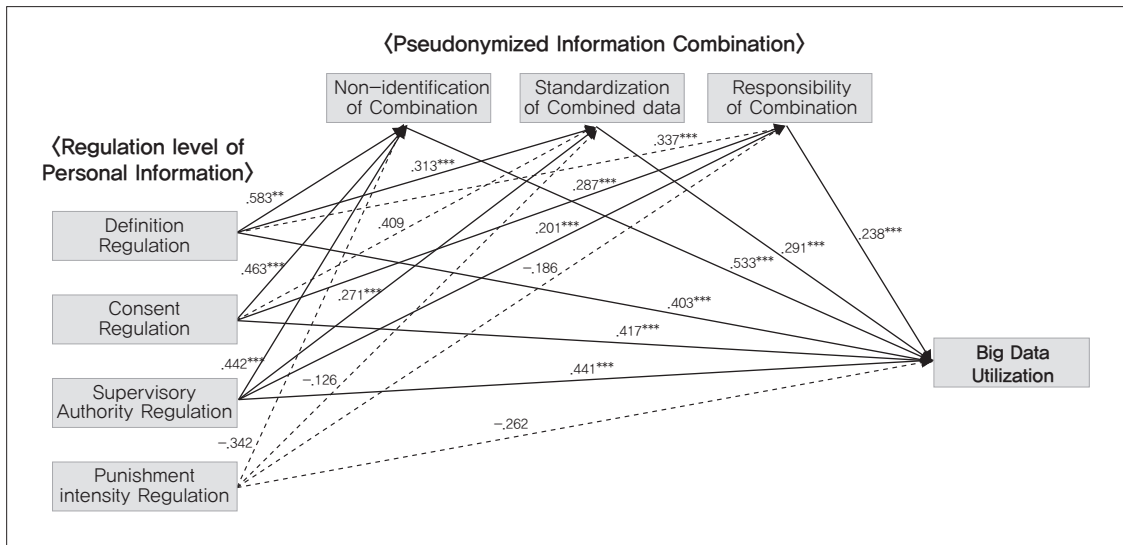
가설검증 결과, 첫째, 개인정보 규제요인과 빅데이터 활용의 관계에 통계적으로 정(+)의 유의한 영향을 미치므로 <가설 1>은 채택되었다. 여기에서 규제강도가 낮을수록(정의규제가 명확할수록, 동의규제가 용이할수록, 감독기구 규제가 통합될수록, 법령위반 처벌강도 규제가 적정수준 일수록) 빅데이터 활용은 증가한다는 것을 의미한다. 다만, 법령위반 적정 처벌강도는 부(-)의 영향을 나타냈다. 다만, 법령위반 시 적정 처벌강도는 빅데이터 활용과의 관계에서 부(-)의 관계를 나타냈으며, 오히려 처벌강도의 강화가 필요함을 시사한다. 즉 현행 개인정보 처벌체계가 형사별 위주이므로 경제별 위주로 개선하되, 법원의 판결 등 현장에서 인식하는 처벌강도가 낮은 '솜방망이 처벌'이 이루어지고 있으므로 법령위반시 일벌백계의 엄격한 법 적용이 필요하다고 할 수 있다.

둘째, 가명정보 결합이 빅데이터 활용에 미치는 영향은 검증하였다. 비표준화 계수는 .553이며, 표준오

차는 .070, C.R(Critical Ratio)는 7.867로서 유의미한 것으로 나타났다(p<.001에서 유의미함). 가명정보 결합요인이 빅데이터 활용과의 관계에서 정(+)의 유의한 영향을 미치므로 <가설 2>는 채택되었다.

셋째, 가명정보 결합의 매개효과를 검증하였다. 먼저 개인정보 규제요인이 데이터 결합에 미치는 영향은 비표준화계수는 .510이며, 표준오차는 .068, C.R(Critical Ratio)는 7.513으로서 유의미한 것으로 나타났다(p<.001에서 유의미함). 개인정보 규제요인이 빅데이터 활용에 미치는 영향으로서 비표준화 계수는 .285이며, 표준오차는 .061, C.R(Critical Ratio)는 4.708로서 유의미한 것으로 나타났다(p<.001에서 유의미함).

<표 10>과 같이, 개인정보 규제요인과 빅데이터 활용과의 관계에서 가명정보 결합의 매개효과는 .118**로 나타났다. 부트스트래핑(Bias Corrected Bootstrapping) 검증방식을 활용하여 가명정보 결합의 매개효과의 유의성을 검증한 결과, 95% 신뢰구간에서 0.18~1.83의 상한값과 하한값을 보이고 있



<그림 2> 경로분석 결과
<Fig. 2> Result of Path Analysis

〈표 10〉 매개효과 검증
 〈Table 10〉 Test of Mediating Effect

Path	Total Effect	Direct Effect	In-direct Effect	Bootstrap Significant(95%)
				Lower ~ Upper
Regulation level of Personal Information → Pseudonymized Information Combination → Big data Utilization	.567	.449	.118**	0.18 ~ 1.83
Regulation level of Personal Information → Non-identification of Combination → Big data Utilization	.567	.356	.211**	1.34 ~ 3.76
Regulation level of Personal Information → Standardization of Combined Pseudonymized Information → Big data Utilization	.567	.503	.064	-1.06 ~ 2.09
Regulation level of Personal Information → Respoinsibility of Combination → Big data Utilization	.567	.532	.035**	0.65 ~ 2.75

* p<.05, ** p<.01

어 “0”을 포함하지 않아 통계적으로 유의하였다. 또한 개인정보 규제요인이 빅데이터 활용의 영향력에 대한 직접효과도 통계적으로 유의하므로, 가명정보 결합은 개인정보 규제요인이 빅데이터 활용에 미치는 영향을 부분매개(Partial Mediation)하는 것으로 나타났다. 세부적으로 매개변수인 가명정보 결합의 구성요인인 결합 비식별성, 결합 가명정보 표준화, 결합 책임성의 매개효과도 살펴보았다. 결합 비식별성은 개인정보 규제요인이 빅데이터 활용에 총효과 .567, 직접효과 .356, 간접효과 .211를 주고 있으며 간접효과의 유의성도 Bootstrapping 검증결과 유의미하였다. 결합 가명정보 표준화는 개인정보 규제요인이 빅데이터 활용에 총효과 .567, 직접효과 .503, 간접효과 .064을 주고 있으며, Bootstrapping 검증결과, 신뢰구간에서 -1.06~2.09의 상한값과 하한값에서 “0”이 포함되어 있어 간접효과의 유의성은 나타나지 않았다. 결합 책임성은 개인정보 규제요인이 빅데이터 활용에 총효과 .567, 직접효과 .532, 간접효과 .035를 주고 있으며, Bootstrapping 검증결과, 간접효과의 유의성도 확인되었다. 그러므로 개인정보 규제요인과 빅데이터 활용과의 관계에서 가명정보 결합이 정(+의) 매개효과

를 갖는다는 〈가설 3〉은 채택되었다. 즉, 개인정보 규제의 명확성은 전반적으로 가명정보 결합의 비식별성, 결합의 책임성이 높을수록 빅데이터 활용에 긍정적인 영향을 미친다고 볼 수 있다. 다만, 부분적으로 결합 비식별성(처벌강도 규제), 결합 가명정보 표준화(동의규제, 처벌강도 규제)와 결합 책임성(정의규제, 처벌강도 규제)에서 매개효과가 나타나지 않았다. 지금까지 데이터 저장과 보호의 관점에서 결합데이터의 표준화가 관리되어 왔다면 향후에는 데이터 관리의 방향을 데이터 활용에 맞추어 바꿀 필요가 있다. 한편, 개인정보 정의의 명확성과 법령위반 적정 처벌강도는 결합 가명정보의 안전성 조치와 재식별 금지 등 결합 책임성에 관계없이 법적의무를 규정하고 있으므로 준수되어야 한다는 것을 시사한다.

4. 조절효과 분석

다음은 개인정보 규제요인의 빅데이터 활용과의 관계에서 결합기관 유형의 조절효과를 분석하였다. 결합기관 유형은 가명정보 결합에 대한 국가의 개입 정도와 기관의 자율성에 따라 자유형, 중개형, 지정형으

로 구분하였다. 결합기관 유형은 비메트릭(범주형) 변수이므로 조절효과 분석은 상호작용효과 분석을 수행하는 메트릭(연속형) 변수와 달리 대응별 모수비교(Pairwise parameter comparison)와 차이 검증을 사용한다. 본 연구에서는 일반적으로 조절변수가 3집단 이상인 경우에 모수의 차이가 0.01 유의수준에서 z 통계량이 ±1.56이 넘으면 가설을 채택하는 대응별 모수비교방법을 사용한다(Bae, 2018). <표 11>과 같이,

결합기관 유형 중 자유형, 중개형, 지정형 순서대로 개인정보 규제요인과 빅데이터 활용과의 관계에서 정(+)의 조절효과가 나타나지 않아 <가설 4>는 기각되었다. 다만, 지정형은 개인정보 동의규제의 빅데이터 활용효과에서 자유형과의 통계량 차이가 -2.450**, 중개형과 통계량 차이가 -2.399**으로 ±1.56보다 큰 값을 보여 결합기관 유형 중 지정형, 중개형, 자유형 순으로 부(-)의 조절효과가 있는 것으로 나타났다.³⁾ ‘분석결과, 자유

〈표 11〉 결합기관 유형의 모수차이 검증
 (Table 11) Pairwise Parameter Comparison Test of Type of Institutes Combining Data

Model	NPAR	CMIN	DF	p	CMIN/DF
Default model	90	116.577	76	.002	1.554
Definition Regulation	Free-type	Intermediary-type	Degsinated-type		
Free-type	.000				
Intermediary-type	-.828	.000			
Degsinated-type	-.475	.391		.000	
Consent Regulation	Free-type	Intermediary-type	Degsinated-type		
Free-type	.000				
Intermediary-type	-.216	.000			
Degsinated-type	-2.450**	-2.399**		.000	
Supervisory Authority Regulation	Free-type	Intermediary-type	Degsinated-type		
Free-type	.000				
Intermediary-type	-.077	.000			
Degsinated-type	-.078	-.002		.000	
Punishment Intensity Regulation	Free-type	Intermediary-type	Degsinated-type		
Free-type	.000				
Intermediary-type	.535	.000			
Degsinated-type	1.237	.828		.000	

* p<.05, ** p<.01

3) 대응별 모수비교(Pairwise parameter comparison)는 두 모수차이를 z통계량을 통해 알 수 있다. 즉 경로계수(C.R.)의 차이값은 각 집단의 회귀계수의 행렬의 임계값이 정(+)의 매개효과 가설이므로 단측검정의 임계값인 ±1.56보다 크면 ‘자유형의 경로계수가 중개형, 지정형보다 작다’는 귀무가설을 기각한다. 이와 같이 3집단 이상의 조절효과 분석의 경우에는 z² 차이 검증은 집단별 모수의 차이가 나타나지 않으므로 대응별 모수비교가 바람직하다(Bae, 2018).

형, 중개형, 지정형 순으로 개인정보 규제요인(정의규제, 감독기관 규제, 처벌강도 규제)에 따라 빅데이터 활용이 증가할 것이라는 결합기관 유형의 조절효과가 각각되었다. 그러나 개인정보 규제요인 중 동의규제에 대해서는 지정형, 중개형, 자유형 순으로 조절효과가 있는 것으로 나타났다. 이러한 결과는 설문대상으로 민간 기업의 비중이 낮아서 기업자율성이 높은 자유형의 조절효과가 낮은 점, 동의규제가 엄격한 상황에서는 국가 개입이 강한 지정형이 오히려 정보주체의 순응에 유리한 점, 그리고 가명정보 결합의 초기에는 개인정보의 재식별 방지를 위한 규제조치로서 결합 전문기관의 지

정과 취소, 결합 가명정보의 안전조치 등에 국가기관의 적극개입을 선호하는 것으로 해석된다.

5. 가설검증 결과

가설검증 결과는 <표 12>와 같다. 개인정보 규제요인과 가명정보 결합은 빅데이터 활용에 정(+의 영향을 가진다. 또한 가명정보 결합은 개인정보 규제요인과 빅데이터 활용의 관계에서 전반적으로 정(+의 매개효과를 가지지만, 결합기관 유형의 조절효과는 나타나지 않았다.

〈표 12〉 가설검정 결과
〈Table 12〉 Hypothesis Test Results

Research Hypothesis	Detailed Hypothesis	Test Results
Hypothesis 1: Regulation factors of personal information will have positive(+) effect on big data utilization.	Definition Regulation → Big Data Utilization(+)	Adoption
	Consent Regulation → Big Data Utilization(+)	Adoption
	Supervisory Authority Regulation → Big Data Utilization(+)	Adoption
	Punishment Regulation → Big Data Utilization(+)	Rejection
Hypothesis 2: Pseudonymized information combination will have positive(+) effect on big data utilization.	Non- Non-identification → Big Data Utilization(+)	Adoption
	Standardization of Combined Pseudonymized Information → Big Data Utilization(+)	Adoption
	Responsibility → Big Data Utilization(+)	Adoption
Hypothesis 3(Mediating Effect): Pseudonymized information combination will have positive(+) mediating effect in relationship between regulation factors of personal information and big data utilization.	Regulation Factors → Non-identification(+) → Big Data Utilization	Partial Adoption
	Regulation Factors → Standardization of Combined Pseudonymized Information(+) → Big Data Utilization	Partial Adoption
	Regulation Factors → Responsibility (+) → Big Data Utilization	Partial Adoption
Hypothesis 4(Moderating Effect): Type of institutes combining pseudonymized information will have moderating effect in order of free-type, intermediary-type, and designated-type in relationship between regulation factors of personal information and big data utilization.	Definition Regulation × Type of Combination Institution → Big Data Utilization	Rejection
	Consent Regulation × Type of Combination Institution → Big Data Utilization	Rejection
	Supervisory Authority Regulation × Type of Combination Institution → Big Data Utilization	Rejection
	Punishment Regulation × Type of Combination Institution → Big Data Utilization	Rejection

V. 결론 및 정책적 시사점

본 연구의 분석결과는 다음과 같다. 첫째, 본 연구에서 개인정보 규제요인은 개인정보 정의 명확성, 개인정보 동의 용이성, 개인정보 감독기구 통합성, 법령위반 적정 처벌강도를, 가명 정보 결합 요인은 결합 비식별성, 결합 가명정보 표준화, 결합 책임성, 결합기관 유형과 통계적으로 관련성이 높은 것으로 나타났다(Kim & Kim, 2020b).

둘째, 개인정보 규제요인은 빅데이터 활용과 정(+)의 유의한 관계를 나타냈다. 즉 개인정보 정의가 명확할수록, 동의요건이 용이할수록, 감독기구가 통합될수록 빅데이터 활용이 증가한다. 다만 법령위반 처벌강도는 빅데이터 활용과 부(-)의 인과관계를 보여 가설이 기각되었다.

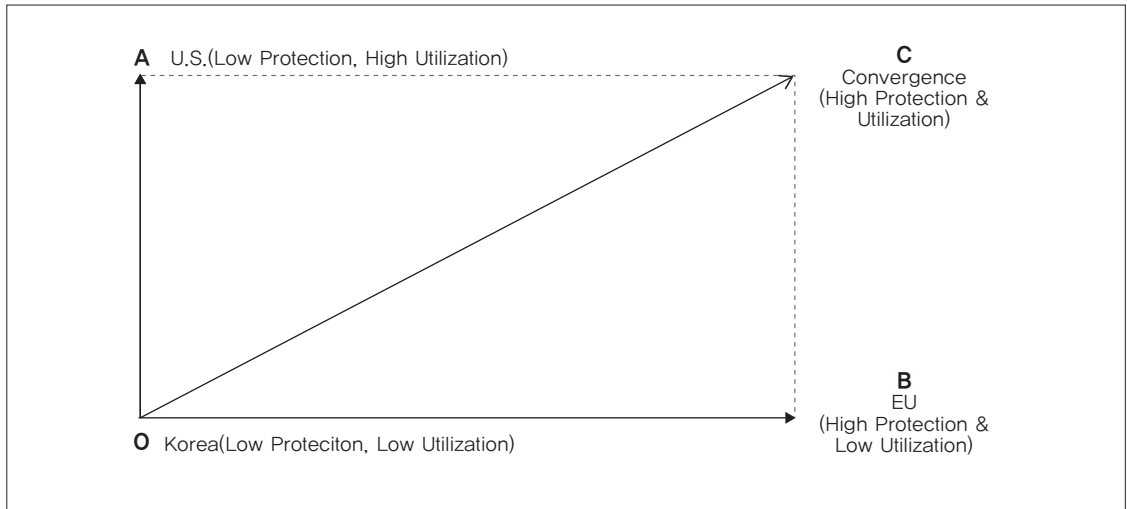
셋째, 가명정보 결합의 매개효과 분석결과, 일부요인은 정(+)의 매개효과가 검증되었다. 구체적으로 가명정보 결합 비식별성 중 정의 규제, 동의 규제, 감독기구 규제는 매개효과가 있으나, 처벌강도 규제는 매개효과가 나타나지 않았다. 그리고 결합 가명정보 표준화 중 정의 규제, 감독기구 규제는 매개효과가 있으나, 동의규제와 처벌강도 규제는 매개효과가 기각되었다. 또한 결합책임성 중 동의규제, 감독기구 규제는 매개효과가 검증되었으나, 정의규제와 처벌강도 규제는 매개효과가 나타나지 않았다. 이러한 결과는 지금까지 데이터 저장과 보호의 관점에서 결합데이터의 표준화가 관리되어 왔다면 향후에는 데이터 관리의 방향을 활용관점으로 전환해야 할 필요가 있음을 시사한다. 한편, 개인정보 정의 명확성과 법령위반 적정 처벌강도는 결합 가명정보의 안전성 조치와 재식별 금지 등 결합 책임성에 관계없이 법적의무를 규정하고 있으므로 준수되어야 한다는 것을 나타낸다.

넷째, 가명정보 결합기관 유형의 조절효과 분석결과,

자유형, 중개형, 지정형 순으로 개인정보 규제요인에 따라 빅데이터 활용이 증가할 것이라는 조절효과가 기각되었다. 그러나 개인정보 규제요인 중 동의규제에 대해서는 지정형, 중개형, 자유형 순으로 부(-)의 조절효과가 있는 것으로 나타났다. 이러한 결과는 설문대상으로 민간기업의 비중이 낮아서 기업자율성이 높은 자유형의 조절효과가 낮은 점, 동의규제가 엄격한 상황에서는 국가개입이 강한 지정형이 오히려 정보주체의 순응에 유리한 점, 그리고 가명정보 결합의 초기에는 개인정보의 재식별 방지를 위한 가명정보의 안전조치가 중요하므로 국가기관의 적극개입을 선호하는 것으로 해석이 된다.

본 연구의 정책적 시사점을 도출하기 위한 이념형(Ideal Type)으로 '착한규제 모형(Good Regulation Model)'을 도출하였다. 본 연구는 개인정보보호와 빅데이터 활용을 조화관계로 파악하며 현행의 낮은 개인정보보호와 낮은 빅데이터 활용형(O형)⁴⁾에서 미국과 유럽의 장점을 절충한 높은 보호, 높은 활용(C형)의 착한규제를 지향한다(Choi, 2017). '착한규제'는 기본적으로 낮은 개인정보보호와 낮은 빅데이터 활용을 trade-off 없이 상호보완(Complementary Relationship)관계로 인식한다. 구체적으로 우리나라는 2015년 신용카드 3사의 1억여건 개인정보 유출 등 2007년부터 2017년까지 개인정보 유출사고는 약 60억 건에 달하여 개인정보 침해가 일상화 되어 개인정보 보호에 취약하다(People Power 21, 2018). 그동안 잦은 개인정보 유출로 개인정보 규제는 강화되었으나, 법원의 솜방망이 처벌로 개인정보 유출기관의 고의·중과실이 명백히 입증되지 않는 한 무죄 또는 가벼운 벌금에 처해졌다. 개인정보 유출사고에 대한 벌금은 미국(2019년 페이스북에 약 6조 원 벌금), EU(2019년 영국항공사에 2,700억 원 벌금)에 비하여 매우 낮은 수준(국내 1위는 2016년 인터파크 약 45억 원)으로 사회적 경각심이 낮아 개인정보 유출사고가

4) 우리나라는 신용카드 개인정보 유출 등 대규모 유출사고가 지속되고 있어 개인정보 규제는 강화되었으나, 과도한 개인정보 사전동의와 모호한 개인정보 정의로 규제의 실효성이 낮아지는 낮은 개인정보보호와 낮은 빅데이터 활용에 직면하고 있다고 비판한다(Lee, 2019; Lee, 2019; Koo,



〈그림 3〉 착한규제 모형
〈Fig. 3〉 Good Regulation Model

반복되는 악순환이 거듭되고 있다. 한편으로 과도한 개인정보 사전 동의와 모호한 개인정보 정의로 규제의 실효성이 낮아 빅데이터 활용도 저조한 규제의 역설에 직면하고 있다(Lee, 2019; Lee, 2019; Koo, 2018).

그러므로 가설검증 결과를 분석하고 착한규제 모형에 입각하여 〈표 13〉과 같이 정책적 시사점을 도출하였다. 총론으로 개인정보보호법 목적조항에 개인정보 보호와 빅데이터 활용원칙 명시하였다. 개인정보 규제요인이 빅데이터 활용에 정(+)의 영향을 미친다는 가설검증결과에 따라 개인정보 정의규제(가명정보 활용범위 구체화), 동의 규제(포괄적 동의 및 사후동의 배제 도입), 감독기구 규제(개인정보보호위원회와 금융위원회의 통합), 처벌강도 규제(형사벌의 경제벌 전환)을 제시하였다. 그리고 가명정보 결합의 매개효과 검증결과에 따라 결합 비식별성(비식별 기술개발 및 인력양성), 결합 가명정보 표준화(결합 전문기관 등 창업지원과 협업체계 구축), 결합 책임성(결합 규제프리존 설정)을 제안하였다.

① 개인정보 보호와 활용원칙 적용, 개인정보위원회로 명칭변경 및 보호와 활용의 전문가로 위원 위촉

개인정보 보호와 활용원칙을 목적조항, 기관명칭, 위원구성에 적용하였다. 개인정보보호법 제1조 목적조항에 개인정보보호와 활용을 명시할 필요가 있다. 더불어 현행 ‘개인정보보호위원회’ 명칭을 ‘개인정보위원회’로 변경하고, 위원구성에 개인정보보호 뿐만 아니라 빅데이터 활용에 정통한 전문가를 위원으로 위촉하여야 한다.

② 개인정보와 가명정보의 활용범위 구체화

개인정보와 가명정보의 활용범위를 구체화하여야 한다. 즉 온라인 식별자로서 IP주소, MAC주소, 쿠키 ID 등을 개인정보 규제에 포함시킬 것인지 명확하게 제시하여야 한다. 또한 민감 정보에 해당하는 의료정보의 경우에는 개인정보의 활용범위를 구체화하여야 한다. 아울러, 신용정보법령에서 규정한 가명정보 결합요건과 비교하여 ‘가명정보 결합 및 반출 등에 관한 고시’를 일치시킬 필요가 있다. 즉 결합전문기관의 O2O(Online to Offline) 온라인-오프라인 공간에서 가명정보 결합여부, 결합 가명정보의 외부반출 승인여부, 외부반출시 가명정보 또는 익명정보 반출기준 등을 명확히 제시하여야 한다.

〈표 13〉 착한규제에 입각한 정책적 시사점
〈Table 13〉 Policy Implications in Good Regulation Model

		Good Regulation Model
Personal Information Regulation Policy	General	<ul style="list-style-type: none"> • Specification of personal information protection and utilization principles, name change to the 'Personal Information Committee' and invitation of experts for personal information protection and utilization
	Regulation Factor of Personal Information	<ul style="list-style-type: none"> • (Definition) Refine the scope of use of pseudonymized information • (Consent) Introduce comprehensive consent and opt-out consent • (Supervisory Authority) Integration of 'Personal Information Protection Committee' with 'Financial Services Commission' • (Punishment Intensity) Converting excessive criminal punishment into economic punishment, including fines
	Pseudonymized Information Combination	<ul style="list-style-type: none"> • (Non-identification) Non-identification technology development and human resource development • (Standardization) Establishment of a collaboration system with other agencies supporting start-ups, through institutes combining personal information. • (Responsibility) Setting of regulation-free zones of pseudonymized information combination

③ 포괄동의 및 사후동의 배제(Opt-out) 도입

포괄동의 및 사후동의 배제 제도를 도입하여 엄격한 개인정보의 사전 동의 제도를 사후규제로 전환하여 위험기반 접근에 맞게 정비하여야 한다. Lee(2019)의 연구결과에서 보듯이, 사전동의시 정보주체가 동의를 확인하고 동의의 의사표시를 하는 비율이 23.2%에 불과하므로 현행의 매건 별 정보주체의 사전 동의를 포괄적 동의 또는 정보주체의 요구 시 사후동의 배제(Opt-out)의 도입을 검토할 필요가 있다.

④ 개인정보보호위원회와 금융위원회 통합

개인정보보호위원회와 금융기관 감독기구(금융위원회)와의 통합이 필요하다. 최근 EU 적합성 평가에서도 한국의 감독기구에서 금융기관이 배제되는 것에 대하여 강하게 비판적 의견을 제시한 것을 보면, 금융기관의 특수성을 인정하는 것은 국제적 추세와도 부합하지 않는다.

⑤ 법령위반 시 과도한 형사처벌을 경제별로 전환

법령위반 시 과도한 형사벌을 경제별로 전환하여야 한다. 개인정보 법령위반에 대한 처벌규제의 개선으로 과도한 형사벌 위주의 처벌에서 과징금, 과태료 등 경

제적 처벌로 전환하는 제도개선이 필요하다.

⑥ 가명정보 결합 비식별 기술개발과 인력양성

가명정보 결합 비식별 기술개발과 인력양성이 필요하다. 데이터 결합의 촉진과 빅데이터 활성화를 위해 차분프라이버시, 동형암호화 등 비식별화 기술의 개발 등 데이터 결합을 적극 지원해야 한다. 결합 가명정보의 표준화는 데이터 관리의 지향점을 기존의 데이터 저장과 보호의 관점에서 앞으로 데이터 활용의 방향으로 전환해야 함을 시사한다.

⑦ 가명정보 결합 전문기관 등 창업지원 협업체계 구축

가명정보 결합기능과 데이터 개방·거래·활용 및 창업지원 기능의 통합연계가 필요하다. 가명정보 결합 전문기관 내부에 데이터 결합 지원센터를 지정하여, 이중 간 데이터 결합의 수요와 공급을 연결해주는 마중물 역할을 수행해야 한다. 특히, 데이터 창업의 아이디어가 있으나, 대기업과 달리 데이터가 부족한 스타트업의 초기 창업지원을 적극 지원해야 한다. 지금까지 데이터 개방기관과 데이터 거래소, 빅데이터 플랫폼 그리고 가명정보 결합 전문기관이 일치하지 않고

관리주체도 달라서 데이터 경제의 시너지 효과가 어려웠다. 그러므로 새로운 서비스와 제품의 개발 등 데이터 스타트업(Data Startup)의 창업지원을 위해 데이터 개방, 가명정보 결합, 데이터 거래소, 빅데이터 플랫폼 센터, 마이데이터 증개기관을 유기적으로 결합하여야 한다.

⑧ 가명정보 결합규제 프리존(Free-Zone) 설정

가명정보 결합규제 프리존(Fee-Zone)을 설정하여 신산업 규제실험의 효과를 검증할 필요가 있다. 프리존에서는 가명정보의 수요자와 공급자를 연결하는 중개기능, 결합 가명정보의 정제 등 품질관리 기능, 창업지원 등 데이터 지원기관과의 협업기능, 결합기관 내부와 외부의 O2O(온라인-오프라인)에서 가명정보 결합이 가능하고 외부반출이 가능한 결합기능을 수행하여야 한다.

본 연구의 한계는 첫째, 설문대상으로 '개인정보보호 전문교육'에 참여한 공공기관과 기업인 등 273명을 선정하여 설문한 결과, 표본의 대표성 또는 표본수를 충분히 확보하지 못한 한계가 있다. 둘째, 연구결과의 일반화를 위하여 설문조사 외에 국제비교연구 2차 통계자료 등 객관적 자료를 대상으로 하는 추가연구가 필요하다. 셋째, 데이터 결합의 조절효과에 대한 분석이 미흡하였다. 결합기관 유형의 조절효과는 가설이 기각되었다. 이러한 결과는 공공부문 종사자가 설문대상자의 다수이고, 교육과 경험 등을 통해 데이터 결합을 아직 경험하지 못한 대상자가 설문에 참여한 결과가 반영된 것으로 보인다. 결합기관 유형은 향후 데이터 브로커의 육성과도 밀접한 관계가 있으므로 추가연구가 필요하다.

향후 연구과제는 첫째, 개인정보자기결정권과 데이터 소유권, 데이터 이동권, 프로파일링 거부권 등 개인정보 보호와 활용에 대한 헌법적 근거와 법률적 규정 사항 등 토의를 통해 세련된 이론 정립과 생산적인 정책생안이 필요하다. 예를 들어, 가명정보의 결합에 의해 생성된 결합 가명정보의 소유권은 누가 갖는가? 개인정보처리자 중 공급자인가, 수요자인가, 아니면 정

보주체인가? 데이터 활용의 법적기초로서 최근에 데이터를 物權에 포함시킬 것이지에 대한 민법상의 논의는 시사하는 바가 크다. 또한 변수에 대한 인식도 조사를 넘어서 보다 객관화되고 일반화된 사례연구 등 질적인 연구가 추가되어야 한다. 둘째, 개인정보 규제요인과 가명정보 결합요인으로 상위에 선정된 요인 중 최근 개인정보보호법 개정안에 포함되지 않은 사항에 대한 법령개정을 위한 후속연구가 필요하다. 지금까지 선행연구에서 제도개선이 필요하다고 지적한 과제에 대한 법령 개정방안이 마련되어야 한다. 구체적으로 개인정보 수집이용 동의요건과 제공동의요건의 일치(Kim, 2017; Lee, 2018), 사후동의 배제(Opt-out) 도입(Kwon, 2015; Cho, 2017), 법령위반시 과다한 형사벌의 경제별로의 전환(Shim, 2015; Kim, 2019), 원칙금지의 포지티브 입법방식에서 원칙허용의 네거티브 입법으로의 전환(Lee, 2019; Choi, et al., 2018) 등이 과제이다(Kim & Kim, 2020b). 셋째, 데이터 개방, 결합 가명정보 표준화 등 개인정보 규제요인이 빅데이터 활용에 영향을 미치는 다양한 조건부 효과를 검토할 필요가 있다. 향후에는 개인정보 규제에 따라 데이터 개방 또는 정보공개의 매개효과 또는 조절효과에 대한 심도 깊은 연구가 필요하다. 넷째, 본 연구는 기기의 Endpoint에서 생성되는 개인정보가 포함된 데이터는 결국 빅데이터가 동일하다는 가정하에 개인정보와 데이터 그리고 빅데이터 개념간의 차이에 주목하지 않았다(Cho, 2017). 그러나 GPS 위치데이터, 기상데이터, 전력데이터 등 개인정보가 필요하지 않은 빅데이터의 활용방안에 대해 개인정보 포함여부, 사전동의 예외가능성, 감독기구의 규제범위, 처벌강도의 적정성 등 분석의 결과가 달라질 수 있으므로, 규제요인과 관련하여 가명정보 결합과 빅데이터 활용에 미치는 영향을 검토할 필요가 있다(Kim & Kim, 2020a).

■ References

- 4th Industrial Revolution Committee (2019). *Recommendation for Government on 4th Industrial Revolution*.
- {4차산업혁명위원회 (2019). <4차산업혁명 대정부 권고안>}
- Bae, B. (2018). *Mediating Effect and Moderating Effect Analysis*.
- {배병렬 (2018). <매개효과 및 조절효과 분석>}
- Barron, R., & Kenny, D. (1986). "The moderator-mediator variable distinction in social psychology research: Conceptual, strategic, and statistical consideration." *Journal of Personality and Social Psychology*, 51(6): 1173-1182.
- Cha, Y. (2019). "Data Economy and the Trends of De-identification Technology in Personal Information Protection." *Weekly Technology Trends*. IITP(Institute of Information & Communications Technology Planning & Evaluation).
- {차연철 (2019). 데이터 경제와 개인정보 비식별기술 동향. <주간기술동향>. 7.10., 정보통신기획평가원.}
- Chae, S. (1997). *Marketing Research*. Seoul: Hakhyeonsa.
- {채서일 (1997). <마케팅 조사론>. 서울:학현사.}
- Chan, F., Chong, A. & Zhou, L. (2013). "An empirical investigation of factors affecting e-collaboration diffusion in SMEs." *International Journal of Production Economics*, 138(2): 329-344.
- Cho, S. (2017). *Improvement Stragy of Information Usability under Personal Information Protection Legal System*. Premium Report of Korea Information Society Development Institute.
- {조성은 (2017). <개인정보보호 법제하에서 정보활용성 향상 전략>. KISDI Premium Report.}
- Choi, H. & Choi, Y. & Lee, S. (2018). *A Study on Improvement of Personal Information Protection System*. Booklet of Sustainable Growth Initiative.
- {최현우·최윤정·이성호 (2018). <개인정보보호제도 개선방안 연구>. 지속성장이니셔티브 발표자료집.}
- Choi, K. (2017). "Artificial Intelligence, Data, and Law" (ed.) *Data Economy*. Seoul: Hans media.
- {최경진 (2017). "인공지능과 데이터, 그리고 법". <데이터 이코노미>, 서울: 한스미디어.}
- Gartner (2011). *How to Plan, Participate and Prosper in the Data Economy*.
- Heo, S. (2014). "Big Data Legal Issues in Korea." *Journal of Law & Public Utilities*, 7(2), 7-21.
- {허성욱 (2014). 한국에서 빅데이터를 둘러싼 법적 쟁점과 제도적 과제. <경제규제와 법>, 7권 2호, 7-21.}
- IDC. (2017). *Worldwide Semiannual Big Data and Analytics Spending Guide*.
- IDC. (2019a). *Worldwide Big Data and Analytics Software Forecast, 2019-2013*.
- IDC. (2019b). *Korea Big data and Analytics Forecast: 2018-2022*.
- IMD. (2018). *IMD World Digital Competitiveness Ranking 2018*.
- Kang, B. (1997). *Multivariate Statistical Analysis using Matrices and SPSS/PC*. Seoul: Hakhyeonsa.
- {강병서 (1997). <행렬과 SPSS/PC 이용 다변량통계분석>. 서울: 학현사.}
- Kang, J. (2017). "An Analysis on the Factors Affecting of Bigdata Utilization in Local Government: Focused on the Public Servants Perception in Seoul." *Journal of Governance Studies*, 12(1), 161-197.
- {강정목 (2017). 지방정부의 빅데이터 활용에 관한 영향요인 분석: 서울시 공무원의 인식을 중심으로. <국정관리 연구>, 12권 1호, 161-197.}
- Kim, I. (2017). "A Study on the Role and Function of Personal Information Protection Commission." *Study on the American Constitution*, 28(2), 219-273.
- {김일환 (2017). 현행 개인정보보호법 체계상 감독기구 법제 정비방안에 관한 연구. <미국헌법연구>, 28권 2호, 219-273.}
- Kim, J. (2017). "Review the Regulation and the Improvement in the Big Data: Focusing on the Policy Directions and Legislation of the New Government." *IT & Law Review*, 15, 157-191.
- {김진영 (2017). 빅데이터 분야 규제와 개선방안 검토: 신정부의 정책방향과 입법과제 중 심으로. <IT와 법 연구>, 15권, 157-191.}
- Kim, J. & Song, T. (2014). "A Study on Initial

- Characterization of Big Data Technology Acceptance: Moderating Role of Technology User & Technology Utilizer.” *Journal of the Korea Contents Association*, 14(9), 538-555.
- {김정선·송태민 (2014). 빅데이터 기술수용의 초기특성 연구: 기술이용자 및 기술활용자 측면의 조절효과를 중심으로. <한국콘텐츠학회논문지>, 14권 9호, 538-555.}
- Kim, S. (2014). “A Study of the Personalization Service and Privacy Paradox in the Big Data Era-Focus on the Sociotechnical Perspective.” *Journal of the Korean Cadestre Information Association*, 16(2), 193-207.
- {김선경 (2014). 빅데이터 시대에 개인화 서비스와 프라이버시 패러독스의 고찰. <한국지적정보학회지>, 16권 2호, 193-207.}
- Kim, S. & Kim, S. (2020a). “The effect of Personal Information Regulation Level and Data Combination on Big Data Utilization.” *Journal of Korea Technology Innovation Society*, 23(2): 306-324.
- {김상광·김선경 (2020a). 개인정보 규제수준과 데이터 결합이 빅데이터 활용에 미치는 영향. <한국기술혁신학회지>, 23권 2호, 306-324.}
- Kim, S. & Kim, S. (2020b). “An Exploration on Personal Information Regulation Factors and Data Combination Factors affecting Big Data Utilization.” *Journal of the Korea Institute of Information Security & Cryptology*, 30(2): 287-304.
- {김상광·김선경 (2020b). 빅데이터 활용에 영향을 미치는 개인정보 규제요인과 데이터 결합요인의 탐색. <정보보호학회논문지>, 30권 2호, 287-304.}
- Kim, T. & Lee, S. (2012). “Regulatory Stringency: with the Application of Report Regulations in Korea.” *Journal of Regulation Studies*, 21(2), 147-185.
- {김태윤·이수아 (2012). 규제강도에 대한 시론적 연구: 신고규제에 대한 적용을 사례로. <규제연구>, 21권 2호, 147-185.}
- Kim, Y. (2019). “A Study on the Improvement of Punishment System for Personal Information Violation.” Doctoral dissertation, Soongsil University.
- {김용학 (2019). <개인정보 침해행위에 대한 처벌제도 개선연구: 형사정책의 지도이념을 통한 검증을 중심으로>. 숭실대학교 박사학위논문.}
- Ko, H. (2019). “A Study on Factors Affecting the Intention to Use Big Data in Businesses.” Doctoral dissertation, Soongsil University.
- {고형석 (2019). <기업의 빅데이터 활용의도에 영향을 미치는 요인에 관한 연구>. 숭실대학교 박사학위논문.}
- Ko, H. (2019). *The Effect on Korea’s Personal Information Protection Law in GDPR era*.
- {고환경 (2019). <GDPR의 한국 개인정보보호 법제에 대한 영향>.}
- Ko, H., Lee, D., Lee, S., Kim, E. & Jeong, J. (2017). *De-identifying Personal Information*. Seoul: PakYoungSa.
- {고학수·이동진·이선구·김은수·정종구 (2017). <개인정보 비식별화 방법론>. 서울: 박영사.}
- Koo, T. (2017). *Regulatory rationalization plan for innovative industry development*. Korean Federation of Science and Technology Societies.
- {구태연 (2017). “혁신산업 발전을 위한 규제합리화 방안”, 데이터사이언스 발전을 위한 규제합리화 포럼 자료집, 한국과학기술단체총연합회.}
- Korean Chamber of Commerce (2018). *Institutional Improvement Survey on Personal Information Protection*.
- {대한상공회의소 (2018). <개인정보보호제도 개선 설문조사>.}
- Kwon, Y. (2015). “Thought on the Self-Determination Right to Personal Information and the Consent Regime.” 2015 Naver Privacy White Paper.
- {권영준 (2015). “개인정보 자기결정권과 동의제도에 대한 고찰.” <2015 네이버 Privacy White Paper>.}
- Lee, E., Oh, B., Im, J., Jang, Y. & Hwang, K. (2017). *A Study on Introduction of Data Linkage & Combination Support System*. Personal Information Protection Commission.
- {이은우·오병일·임진희·장여경·황규만 (2017). <데이터 연계·결합 지원제도 도입방안 연구>. 개인정보보호위원회.}
- Lee, H. (2019). “A study on Legislation of Social Regulation: the Perspective of Regulatory Paradox.” *Journal of Parliamentary Research*, 13(1): 5-69.

- {이혜영 (2019). 규제역설 관점에서 본 사회적 규제 입법의 한계와 과제: 개인정보보호 규제를 중심으로. <의정논총>, 13권 1호, 5-69.
- Lee, J. & Jang, M. (2018). "A Study on the Intention to Use Big data Based on the Technology Organization Environment and Innovation Diffusion Theory in Shipping and Port Organization." *Journal of Korea Port Economic Association*, 34(3), 159-181.
- {이준필·장명희 (2018). TOE와 혁신확산이론에 따른 해운항만조직의 빅데이터 사용의도에 관한 연구. <한국항만경제학회지>, 34권 3호, 159-181.}
- Lee, S. (2016). "Research on Determinants for Big data System Adoption in Organization." Doctoral dissertation, Sungkyunkwan University.
- {이선우 (2016). <조직에서의 빅데이터 시스템 도입을 위한 결정요인에 대한 연구>. 성균관대학교 박사학위논문.}
- Lee, S. (2017). *Suggestion of Improvement Plan for advancement of Personal Informantion Protection Act*. Conference Booklet of 50th Anniversary of E-Government.
- {이성엽 (2017). "개인정보법 선진화를 위한 개선방안 제안." 빅데이터 시대의 개인정보활용과 책임, 전자정부50주년 학술세미나 자료집.}
- Lee, S. (2019). "Personal Information Utilization and Role of Government." (ed.) *Regulatory Reform Task and Case Study for promoting New Industries*. Jinhan M&B.
- {이상무 (2019). "개인정보 활용과 정부의 역할", <신산업촉진을 위한 규제개혁 과제와 사례연구>, 진한 M&B.}
- Lian, J., Yen, D. & Wang, Y. (2014). "An exploratory study to understand the critical factors affecting the decision to adopt cloud computing in Taiwan hospital." *International Journal of Information Management*, 34, 28-36.
- Ministry of Interior and Safety (2020). *Ex-notice on Combination and Export of Psudonymized Informatioin*.
- {행정안전부 (2020). <가명정보의 결합 및 반출 등에 관한 고시(안)>.}
- Oliveira, T., Thomas, M. & Espadanal, M. (2014). "Assessing the determinants of cloud computing adoption: An analysis of the manufacturing and services sectors." *Information and Management*, 51, 497-510.
- Park, K. (2018). "Big data Possibility in Personal Information Protection Legal system." (ed.) *Personal Informaion in the 4th Industrial Revolution Era*. Seoul: Sechang publisher.
- {박귀희 (2016). <행정서비스에서 빅데이터 활용의 결정요인에 관한 연구: 데이터 품질관리를 중심으로>. 계명대학교 박사학위논문.}
- Park, N. (2016). "Improvement Review of Korea's Personal Information Protetion Legal system according to the Analysis of Personal Information Protection Act in major countries on Big Data." <2016 Naver Privacy White Paper.>
- {박노형 (2016). 빅데이터 관련 주요국가의 개인정보보호법 제도 분석에 따른 한국 개인정보보호법 개선의 검토. <2016 네이버 Privacy White Paper.>}
- People Power 21 (2018). *2007-2017 Worst 44 in Personal Information suffering history*. Research paper.
- {참여연대(2018). <2007-2017 개인정보 수난사 Worst 44>. 연구보고서.}
- Personal Information Protection Commission. & Ministry of the Interior and Safety. (2018). *2018 Survey of Personal Information Protection in Korea*.
- {개인정보보호위원회·행정안전부 (2018). <2018 개인정보보호 실태조사>.}
- Shin, Y. (2018). *A Study on Precautionary Regulatory Status Analysis and Improvment Plan for Personal Information Protection in Korea*. Autume Conference Booklet in the Korean Association for Policy Studies, 17-41.
- {신영진 (2018). "우리나라 개인정보보호를 위한 사전예방적 규제현황분석 및 개선방안에 관한 연구." 한국정책학회 추계학술대회자료집, 17-41.}
- Tornatzky, L. G., Fleischer, M., Chakrabarti, A. K. (1990). *Processes of Technological Innovation*. Lexington Books.
- Venkatesh, V., Moris, M. G., Davis, G. B., Davis, F. D. (2003). "User acceptance of information

- technology: Toward a unified view.” *Information Management*, 27(3): 425-478.
- Weigend, A. (2018). *Data for the people: how to make our post-privacy economy work for You*, Trans. Hong, J. Paju: Four Season.
- {안드레아스 와이겐드 저 홍지영 역 (2018). <포스트 프라이버시 경제>, 파주: 사계절.}
- Woo, S. (2018). “A Study on the Factors Affecting Personal Information De-identification Based on Big Data in the Financial Industry.” Doctoral dissertation, Soongsil University.
- {우순규 (2018). <금융산업에서 빅데이터 기반의 개인정보 비식별 조치에 영향을 미치는 요인에 관한 연구>. 숭실대학교 박사학위 논문.}
- Yang, H. (2015). “the Effects of Applying Personal Information De-identification Technology on Intention to use Big data: Big data providers and consumers perspectives.” Doctoral dissertation, Kwangwoon University.
- {양현철 (2015). <개인정보 비식별화 기술 적용수준이 빅데이터 이용활성화 의도에 미치는 영향: 빅데이터 공급자와 수요자 관점>. 광운대학교 박사학위 논문.}
- Yoon, H. (2014). “The effect on Big Data Utilization and Effectiveness of Organizational, Technological, Environmental Factors.” *Research on Business Management*, 7(2): 153-178.
- {윤희정 (2014). 조직적, 기술적, 환경적 요인이 빅데이터 활용 및 효과에 미치는 영향. <경영관리연구>, 7권 2호, 153-178.}
- Yoon, H. (2018). “A Search for a new Regulatory Paradigm for Big data.” *Journal of Law and Economic Regulation*, 11(1), 71-94.
- {윤혜선 (2018). “빅데이터 규제의 새로운 패러다임 모색을 위한 연구.” <경제규제와 법>, 11권 1호, 71-94.}
- Yoon, K. (2018). *A Study on Administrative Service Innovation using Information Technology in 4th Industrial Revolution Era*. Research Paper of ‘Korea Institute of Public Administration’, 2018-16.
- {윤광석 (2018). <4차 산업혁명시대 정보기술을 활용한 행정서비스 혁신방안 연구>. <한국행정연구원 연구보고서>, 2018-16.}
- Yoon, S, & Oh, J. (2018). “A Study on the Factors that Affect the Intention to Use Big Data by IT, Human and Organizational Resources.” *Journal of the Korean Operations Research and Management Science Society*, 43(4), 67-94.
- {윤수영·오재인 (2018). IT자원, 인적자원, 조직적 자원에 따른 빅데이터 사용의도에 미치는 요인에 관한 연구. <한국경영과학회지>, 43권 4호, 67-94.}