

빅데이터 활용에 영향을 미치는 개인정보 규제요인과 데이터 결합요인의 탐색*

김 상 광,^{1*} 김 선 경^{2†}
^{1,2}충남대학교 (대학원생, 교수)

An Exploration on Personal Information Regulation Factors and Data Combination Factors Affecting Big Data Utilization*

Sang-Gwang Kim,^{1*} Sun-Kyung Kim^{2†}
^{1,2}Chungnam National University (Graduate student, Professor)

요 약

그동안 데이터 개방, 기술수용이론 등 빅데이터 활용의 영향요인에 대한 법·정책적 연구는 다수 있었으나, 제약 선 역할을 하는 개인정보 규제요인 또는 데이터 결합요인이 빅데이터 활용에 미치는 영향에 대한 실증연구는 거의 없었다. 이에 본 연구는 델파이 분석(Delphi Analysis)을 통해 빅데이터 활성화에 부정적(-) 관계를 보이는 개인정보 규제요인과 긍정적(+) 효과를 보이는 데이터 결합요인이 무엇으로 구성되는지 요인의 우선순위를 이론적으로 탐색하였다. 델파이 분석결과, 개인정보 규제요인은 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 순으로 상위요인이 조사되었다. 다음으로 데이터 결합요인은 결합 비식별성, 결합데이터 표준화, 결합 책임성, 결합기관 유형, 결합경험, 결합 기술가치 순으로 조사되었다. 이러한 연구결과는 빅데이터 활성화를 위해 개인정보 규제와 데이터 결합정책 설계 시 어느 구성요인을 우선적으로 제도개선 해야 하는지 시사점을 제공한다.

ABSTRACT

There have been a number of legal & policy studies on the affecting factors of big data utilization, but empirical research on the composition factors of personal information regulation or data combination, which acts as a constraint, has been hardly done due to the lack of relevant statistics. Therefore, this study empirically explores the priority of personal information regulation factors and data combination factors that influence big data utilization through Delphi Analysis. As a result of Delphi analysis, personal information regulation factors include in order of the introduction of pseudonymous information, evidence clarity of personal information de-identification, clarity of data combination regulation, clarity of personal information definition, ease of personal information consent, integration of personal information supervisory authority, consistency among personal information protection acts, adequacy punishment intensity in case of violation of law, and proper penalty level when comparing EU GDPR. Next, data combination factors were examined in order of de-identification of data combination, standardization of combined data, responsibility of data combination, type of data combination institute, data combination experience, and technical value of data combination. These findings provide implications for which policy tasks should be prioritized when designing personal information regulations and data combination policies to utilize big data.

Keywords: Personal information, Big data utilization, Regulation factors, Data combination, Delphi analysis

Received(02. 19. 2020), Modified(03. 18. 2020),
Accepted(03. 18. 2020)

* 본 연구는 2018년도 충남대학교 해외과전 연구비 지원에

대한 결과입니다.

† 주저자, tim2co@gmail.com

‡ 교신저자, kimsk@cnu.ac.kr (Corresponding author)

I. 서론

아마존, 구글, 페이스북, 알리바바, 바이두 등 빅테크는 인터넷 시대를 넘어 데이터 경제로 진입하는 시기에 소비자의 선호와 취향을 대표하는 개인정보를 수집·활용하는데 온 힘을 쏟고 있다. IDC(2017)에 따르면 빅데이터는 소비자와 생산자의 끝단에 있는 스마트폰, 카메라, 웨어러블기기, IoT센서에서 나온 Endpoint 데이터의 비중이 약 70%를 차지하여 비중이 가장 높으며, 또한 생성되는 데이터의 대부분이 개인과 연결된 정보이므로 데이터와 개인정보를 동일한 의미로 사용한다(조성은, 2017). 그러나 개인정보가 포함된 데이터는 정보주체의 동의, 암호화 등 안전조치를 거쳐 제3자에게 제공되거나 데이터 결합에 활용되는 등 프라이버시 보호와 데이터 활용은 이해관계가 상충하는 패러독스(Paradox) 성질을 가진다(김선경, 2014; 이상무, 2019). 미국의 비식별정보(De-identification information), EU GDPR의 가명정보(Pseudonymous information), 일본의 익명가공정보(Anonymously processed information) 활용에서 보듯이, 세계 각국은 데이터 활용을 위해 누구의 정보인지 알지 못하도록 개인정보의 식별성을 완화하는 비식별조치, 가명화, 익명화 등 데이터 결합을 촉진하고 있다. 또한 개인정보의 정의 명확화, 동의규정 간소화, 감독기구의 통합, 법령위반시 과징금 강화 등 데이터 경제 활성화와 AI First를 위한 정책경쟁을 치열하게 전개하고 있다.

한편 우리나라의 개인정보 규제수준은 세계적으로 엄격하고, 빅데이터 활용은 저조한 수준이다. 2018년 IMD World Digital Competitiveness Ranking에서 한국은 31위로 빅데이터 활용도 낮고, 2017년 OECD 자료에 의하면 한국의 빅데이터 기업비중은 4%로서 매우 저조한 수준이다. 국제적 평가와 더불어 대기업은 빅데이터 활용의 걸림돌로 개인정보 수집·이용 및 제3자 제공에 대한 사전동의 규제로 인해 빅데이터의 구축·유통·활용이 위축되어 구글 등 빅테크와의 경쟁에서 뒤처진다고 어려움을 호소하고 있다. 또한 스타트업은 전통산업에 비해 결합데이터가 부족하여 신산업 창출이 지연되고 있다고 한다(관계부처 합동, 2019). 특히 결합데이터의 부족은 기업의 빅데이터 활용역량의 저하(2018년 빅데이터 분석경험 6.6%)와 함께 기계학습을 통한 AI 알고리즘의 개발에 부정적 영향을 미친다. 또한 데이

터 결합을 통한 시장조사, 마케팅 전략 개발, R&D, 맞춤형 서비스 등 빅데이터 활용역량은 신산업 서비스 개발의 근간이 되고 있어 기업의 경쟁력에 미치는 영향력이 지대하다.

최근 개인정보와 빅데이터에 관한 연구는 크게 개인정보 보호제도의 문제점과 비식별조치 등 데이터 결합에 대한 연구(김진영, 2017; 조성은 2017; 구태연, 2017; 박노형 2016; 이상무, 2019), 그리고 빅데이터 활용의 영향요인으로 데이터 결합의 기술적·관리적 접근 및 국내의 비교연구(이선우, 2016; 고태수 외, 2017; 우순규, 2018; 양현철, 2015; 이은우 외, 2017)가 있다. 그러나 빅데이터 활성화에 개인정보 규제개혁이 중요함에도 불구하고 개인정보 규제와 데이터 결합에 대한 개념과 구성요소에 대한 실증연구는 거의 이루어지지 않았다.

다행스럽게도 최근 법안상정 14개월만인 2020년 2월에 과학적 연구 등 특정목적하에 정보주체의 동의없는 가명정보의 활용을 허용하고, 부처별로 분산된 감독기구를 개인정보보호위원회로 통합하는 것을 골자로 하는 데이터3법 개정안(개인정보보호법, 정보통신망법, 신용정보법)이 개정되었다. 그러므로 향후 관련 법이 정책적으로 안착하기 위해서는 빅데이터 활용에 특히 중요한 개인정보 규제와 데이터 결합요인에 대한 보다 구체적이고 실증적인 연구가 요구되는 시점이다.

본 연구는 빅데이터 활성화를 위해 개인정보 규제요인과 데이터 결합요인의 필요성에 대한 이론적 근거를 살펴보고, 델파이 분석을 통해 개인정보 규제와 데이터 결합의 구성요인을 시론적으로 탐색한다. 도출된 구성요인의 우선순위를 토대로 향후 빅데이터 활용을 위한 개인정보 규제정책의 설계시 어떤 구성요인의 제도개선에 역점을 두어야 하는지 개선방향과 우선순위에 대한 시사점을 모색하고자 한다.

II. 빅데이터와 개인정보 규제요인

최근 빅데이터 활성화를 위한 개인정보 규제혁신이 주목받고 있다. 2018년 개인정보보호 실태조사에 의하면 공공기관의 빅데이터 분석활용 경험은 15.1%, 민간기업은 6.6%로서 빅데이터 활용경험은 매우 저조하다. 또한 빅데이터 활용의 제약점으로 개인정보 정의의 불명확성, 비식별조치의 법적근거 미흡, 사전동의 규제의 엄격성, 형사벌 위주의 처벌, 규제법률의 중복혼선 등 개인정보 규제수준이 주요한

요인으로 나타나고 있다(조성은, 2017; 최현우 외, 2019). 그러므로 빅데이터 활용을 촉진하는 요인으로 최근에 새롭게 대두되는 개인정보 규제수준과의 연관성을 찾는 일은 매우 의미있는 작업이다.

규제 일반론 관점에서 보면, 개인정보 규제수준이 높을수록 빅데이터 활용이 감소한다는 부정적(-) 의견이 제기되어 왔다(윤혜선, 2018; 최진욱, 2006; 최성락, 2018). 즉 개인정보의 정의와 범위가 불명확할수록, 정보주체의 사전동의 획득이 어려울수록, 감독기구가 중복되어 혼선을 보일수록, 법령위반시 처벌강도가 형사별 위주로서 적정수준보다 높을수록 데이터 결합은 위축되고 결국 빅데이터 활용은 저조하게 된다는 것이다.

지금까지 개인정보 규제수준의 정의와 개념에 대한 체계적인 연구는 거의 없으나, 김태윤·이수아(2012)는 규제일반론에 기초하여 산업안전·환경·기업경쟁 등 영역별 ‘규제강도 결정요인’을 재분류하여 ‘규제강도’를 개념화하였다. 즉 규제강도는 ‘행위완결의 불확실성과 의무 준수부담 및 처벌강도의 종합적 수준’이라고 정의하고 있다. 이러한 규제강도 결정요인을 개인정보 분야에 적용하여 개념화하면, 개인정보 규제수준(Personal information regulation intensity)은 ‘개인정보에 관한 행위의 불확실성과 의무 준수부담 그리고 법령위반시 처벌강도의 종합적 수준’이라고 할 수 있다. 여기에서 의무의 준수부담은 법령위반시 처벌에 따른 위험부담과 일맥상통하므로 본 연구에서는 ‘행위의 불확실성’과

‘의무 준수부담’으로 나누어 구성요인을 고찰한다.

먼저 개인정보 규제요인으로 ‘행위의 불확실성(Uncertainty of behavior)’에 해당하는 요인은 법적으로 요구되는 법률행위가 정의가 불명확하다든지, 행위의 요건이 모호하다든지 하여 의사결정의 지연 등으로 비용이 발생하는 경우이다. <표 1>에서 보는 바와 같이, ‘2018년 개인정보보호 실태조사’와 연구자의 선행연구를 토대로 개인정보 정의 불명확성, 데이터 결합의 근거 모호, 비식별 가이드라인의 법적 근거 미흡, 가명정보 등 활용제도 미도입, 원칙급지의 포지티브(Positive) 입법규제 등이 대표적인 행위 불확실성에 해당하는 요인이다(조성은, 2017; 박노형, 2016; 김진영, 2017).

다음으로 ‘의무 준수부담(Burden of obligatory compliance)’ 기준에 해당하는 요인으로는 대한상공회의소 설문조사와 문헌연구 검토결과, 엄격한 개인정보 사전동의, 법령위반시 형사별 위주의 강한 처벌강도, 개인정보 감독기구의 중복혼선, 개인정보 수집·이용과 제공시 동의 요건의 차이, 개인정보 유출시 엄격한 신고통지제도 운영, 클라우드 서비스 이용시 불합리한 개인정보 위수탁 규제 등이 있다(오정현·선미란, 2015; 이상무, 2019; 김용학, 2019).

본 연구는 개인정보 규제수준 요인의 기준으로 ‘행위의 불확실성’과 ‘의무준수부담’을 설정하고 두 가지 기준에 해당하는 구성요인을 개인정보보호 실태조사와 대한상공회의소 설문조사 등 선행연구에서 추출하여 <표 1>과 같이 16개의 후보요인을 도출하였다.

Table 1. Candidate Factors of Personal Information regulation factors

Variables	Component factors of Personal Information Regulatory Level		Source
	Factors	Attributes	
Uncertainty of behavior	Dispersion of Personal Information Supervisory Authority	The current Personal Information Protection Commission preside over policy Deliberation, Ministry of the interior and Security takes on public institutions, the Korea Communications Commission supervises online, and the Financial Services Commission supervises financial agency	Kim, I. (2017), Cho, S. (2017)
	Ambiguity of Personal Information Definition	Personal information consists of personally identifiable information (identified information) and information that can be personally identifiable and easily combined with other information (identifiable information), so that at the concept of ‘Identifiable Information’ is uncertain and relies on Judicial precedent and Administrative interpretation	Kim, J. (2017), Korean Chamber of Commerce. (2018)
	Non-introduction of Pseudonymous Information etc.	There is no provision for the concept of Pseudonymous Information, Anonymously Processed Information, and De-identification Information that uses personal information safely without consent of the information subject.	Park, N. (2016), Korean Chamber of Commerce. (2018)

Variables	Component factors of Personal Information Regulatory Level		Source
	Factors	Attributes	
Uncertainty of behavior	Lack of Consistency among Personal Information Acts	Similar obligations such as collection & use of personal information, offering to third parties, destruction and consignment of Personal Information are overlapped and distributed among Personal Information Protection, resulting in confusion(Personal Information Protection Act, Information and Communication Network Act and Credit Information Act).	Kim, S. (2015), Korean Chamber of Commerce. (2018)
	Ambiguity of Data Combination Regulation	There is no consensus on Data Set Combination provision, which makes the basis for data combination between institutions controversial	Ko, H. et al (2017), Cho, S. (2017)
	Ambiguity of De-identification requirements	Requirements for De-identification information are defined in the Administrative Guidelines without legal basis, causing the problem of legal effects such as Re-identification of combined data.	Kim, J. (2017), Korea Internet Security Agency (2016)
	Negative legislation of 'Permission Principle, No exception'	Currently, Positive Legislation that prohibits matters that are not regulated in Act, such as new data collection and combined services, is not friendly to fostering new IT industries.	Lee, S. (2019), Korean Chamber of Commerce. (2018)
Burden of Legal Obligatory Compliance	Difficulty of Personal Information Consent	Strict informed consent of information subjects is required except for urgent cases such as 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.	Kwon, Y. (2015), Cho, S. (2017)
	Difference of Consent Requirement between collection & use of Personal Information and Offering to third Party	When providing personal information to third party, requirements are stricter than collection & use of personal information by providing consent of information subject to contracts, or in case of the legitimate interest of the information processor.	Kim, K. (2017), Personal Information Protection Commission. (2018)
	Lack of Self-Regulation in private association	Similar to Japan, a self-regulatory system has not been introduced in which sectoral private associations autonomously inspect and penalize member companies.	Shin, Y. (2018)
	Non-Introduction of Data Portability and Profiling Prohibition of Personal Information	Unlike EU GDPR, Data Portability, the right to move personal data to other provider to enforces data subject's rights, and Prohibition of Profiling, preventing automated collection of personal information, were not introduced in Act.	Kim, B. (2018), Lee, E. (2017)
	Strict Report Notification System in case of Personal Information Leakage	If personal information is leaked, there are heavy burden that the information subject is to be notified, and if more than 1,000 people are leaked, Leaking party is required to report to administrative office even no damages.	Lee, S. (2017)
	Non-Introduction of Data Protection Officer(DPO)	The current Personal Information Protection Officer is able to co-operate with other tasks, so permanent staff(DPO) is not appointed	Lee, S. (2017)
	Irrational Personal Information Entrustment Regulation	When a company entrusts the cloud service, the burden on the company is high because it requires the consent of the information subject.	Korean Chamber of Commerce.(2018)
	Excessive Criminal Punishment in case of violation of law	In case of violation, heavy criminal penalties such as imprisonment, fines are imposed instead of administrative penalties.	Kim, Y.(2019)
	Lower penalty compared to EU GDPR	EU GDPR imposes a penalty of 4% or 20 million euros (approximately 26 billion won) of sales for serious violations, while Korea's Personal Information Protection Act is relatively low due to a fine of up to 500 million won.	Personal Information Protection Commission. (2018)

III. 빅데이터와 데이터 결합요인

3.1 데이터 결합과 빅데이터와의 관계

데이터 결합(Data Combination)은 두 개 이상의 출처로부터 동일인이나 동일한 사건에 연관된 정보를 함께 가져와서 단일 출처의 정보만으로 알기 힘든 정보요소간의 관계를 밝히는 것으로 '데이터의 이종간 연계 및 통합'을 의미한다(이은우 외, 2017). 데이터 결합은 개인정보(원본정보)를 누구의 정보인지 알 수 없도록 개인의 식별성을 제거한 비식별정보, 가명정보로 전환하고 암호화 등 안전조치를 거쳐 이종간 데이터를 연계하는 것으로서, 정보주체의 동의없이 이용하는 빅데이터 활용의 핵심활동이다. 데이터 결합을 통해 정부는 선거, 병역, 조세, 복지 등 공익목적으로 서비스를 제공하고, 기업은 고객정보를 활용한 시장조사, 신약개발, 신상품개발, 마케팅 전략 수립 등 부가가치를 높여 시장에서 경쟁한다. 최근 데이터 3법의 개정으로 데이터 결합을 통해 인공지능(AI) 경제가 2030년까지 455조원의 경제효과를 보일 것으로 전망한다. 이처럼 데이터 결합을 통한 비식별정보의 활용으로 IT, 핀테크, 의료 등 데이터 산업의 발전이 가속화 될 것이다(이상무, 2019).

데이터 결합의 절차는 비식별조치-적정성 평가-사후조치의 단계로 이루어진다. 먼저 비식별조치는 가명처리, 데이터삭제, 범주화 등 비식별기술을 활용하여 개인 식별요소를 제거하고, 적정성평가는 비식별조치가 적정하게 이루어졌는지 외부평가를 통해 평가한다. 이어서 사후조치에서는 비식별정보의 재식별 금지를 위한 안전조치를 취하고 관리하는 활동이다. 여기에서 비식별조치는 데이터결합의 1단계에 해당하는 기술적 조치를 의미한다. '개인정보의 비식별조치'는 데이터 값 삭제, 가명처리, 총계처리, 범주화, 데이터 마스킹 등을 통해 개인정보의 일부 또는 전부를 삭제하거나 대체함으로써 다른 정보와 쉽게 결합하여도 특정 개인을 식별할 수 없도록 하는 일종의 익명화 조치를 말한다(관계부처 합동, 2016; 우순규 2018).

구체적으로, 개인정보를 활용하는 방식은 크게 최초 수집목적과 다른 목적으로 정보주체의 동의 없이 이용하는 '목적 외 이용방식'과 개인정보의 식별성을 제거하여 빅데이터로 이용하는 '비식별화(데이터 결합) 방식'이 있다(박노형·정명현, 2017; 이대희, 2018). 목적 외 이용방식(EU GDPR)은 최초 수집

목적과 합리적인 관련성이 있는 경우 또는 '기록보존·과학연구·통계작성 목적'으로 가명처리한 경우에 정보주체의 추가적인 동의 없이 개인정보를 활용하도록 재량을 허용하는 방식이다. 목적 외 이용방식은 국가별 고유한 법제도에 따라 서로 상이한 반면에, 비식별화 방식은 나라마다 보호와 활용의 공통수단으로서 빅데이터 활용의 대안으로 각광받고 있다.

일반적으로 데이터 결합과 빅데이터 활용의 관계는 수단과 목적관계 또는 긍정적(+) 상관관계를 갖는다. 데이터 결합을 통해 개별 데이터가 보유한 관계성과 영향관계를 추가로 밝혀 이전에 몰랐던 새로운 사실과 가치를 증가시키기 때문이다¹⁾. 따라서, 개인정보가 포함된 데이터의 연계와 결합을 통해 개인정보처리자는 새로운 부가 가치를 창출할 수 있다.

3.2 데이터 결합요인 검토

그간 데이터 결합요인과 관련한 연구는 제약적인 수준으로 이들 관련 연구를 살펴보면 데이터 결합은 크게 기술적 접근과 관리적 접근으로 구분하여 살펴볼 수 있다(고학수 외, 2017; 이은우 외, 2017). 먼저, 데이터 결합의 기술적 접근(Technological approach)은 개인정보 비식별화 기법, 결합데이터 표준화, 결합 기술가치, 결합비용, 결합경험을 들 수 있다. 먼저 '개인정보 비식별화 기법은 개인 식별자를 제거하거나 대체하는 데이터 삭제, 가명처리 등 기술적 조치를 취하는 것과 연결하여 k-익명성 등 프라이버시 모델을 적용하여 다른 정보와 결합가능성을 검토하고 추론을 통해 개인이 재식별되는 위험성을 방지하는 적정성 평가를 거친다. 비식별조치는 개인 식별성을 제거함으로써 개인정보 오남용과 유출 가능성을 획기적으로 낮추고 개인정보에 적용되는 복잡한 법적규율로부터 벗어날 수 있어 데이터 활용을 높일 수 있다(고학수 외, 2017). 결합데이터 표준화는 결합이 용이하도록 데이터 종류, 형식, 구성요소 등 상호호환이 가능하도록 일치시키는 방법이다. 결합의 기술가치와 비용은 결합으로 인한 서비스와 데이터의 가치 등 편익에서 데이터 결합과 가공, 분석에 소요되는 비용을 제외하여 결합의 가치를 산정한다. 데이터 결합경험은 기존에 데이터 결합과 분석을 수행했는지 여부로서 빅데이터 시스템 도입의 중요한 영향

1) 양현철(2015)은 개인정보 비식별화 기술수준이 높을수록 빅데이터 사용의도에 미치는 긍정적 영향을 실증 분석하였다.

요인으로 연구되었다.

한편 최근에는 비식별정보가 어떤 조건하에서 누구에게 어떻게 활용되었는지에 대한 절차적, 관리적 요소를 고려하는 데이터 결합의 관리적 접근(Managerial approach)이 부각되고 있다. 고훈수외(2017)는 데이터 결합의 기술적 접근으로서 비식별조치의 우수성은 인정하더라도 100% 비식별 또는 재식별 위험이 없어지지 않은 한계점을 감안하여, 관리적 체계가 보다 중요하다고 주장한다. 데이터 결합의 관리적 접근은 결합기관 유형(결합 거버넌스), 결합혁신의 필요성, 최고경영층의 결합지원, 결합 수용성, 결합에 대한 경쟁자 압력 등으로 나누어 볼 수 있다.

데이터 결합기관 유형(데이터 결합 거버넌스)은 기업간 데이터결합시 데이터 중개, 결합키 제공, 반출데이터 허용 등을 수행하는 전문기관과 데이터 관리를 누가 어떻게 수행하느냐의 문제로서 거버넌스 유형에 따라 자유형, 중개형, 지정형으로 구분할 수 있다. 결합혁신의 필요성은 조직이 시장개척, 조사연구 등 데이터 결합을 통해 신상품과 서비스를 개발하고자 하는 조직적 요구와 노력정도를 의미한다. 최고경영층의 결합지원은 관리적 측면에서 데이터 결합조직과 예산, 인력 등을 구성하는 간부진의 지원을 말한다. 결합 수용성은 정보주체와 처리자의 데이터 결합에 대한 호의적인 용인도를 의미한다. 결합에 대한 경쟁자 압력은 시장에서 다수의 경쟁자가 등장하여 데이터결합 서비스를 제공하는 경쟁환경을 지칭한다. 이상 <표 2>와 같이, 데이터 결합은 기술적 접근과 관리적 접근으로 구분하고 각각의 하위 구성요소로 정리하였다.

Table 2. Data combination approach

Technological approach	Managerial approach
<ul style="list-style-type: none"> • De-identification Technique • Standardization of Combined Data • Technological value of De-identification • Data combination Cost • Data combination Experience 	<ul style="list-style-type: none"> • Type of Data combination institute(Governance of Data Combination) • Innovation necessity of data combination • Top management support of data combination • Receptivity of data combination • Competitor pressure of data combination

이와 더불어 국가 간 데이터 결합의 정책적 동향을 검토해 봄으로써 국가별 관리적 접근의 공통요소를 검토하였다. 나라마다 데이터 결합에 있어 공통적으로 데이터 결합의 법령개정, 데이터 결합의 책임성, 데이터 결합에 친화적인 환경조성을 위한 결합정책 등에서 관리적 요소를 공유하고 있기 때문에 이를 검토하여 구성요인 도출에 접근하는 것이 의미가 있다.

예를 들어, 미국, EU, 일본 등 선진국은 안전한 데이터 활용을 위해 비식별정보, 가명정보, 익명가공정보 등을 도입하여 데이터 활용을 위한 개인정보의 범위를 확장하고, 결합데이터의 안전성 검사 등 데이터 결합을 위한 법령개정을 추진하고 있다. 또한 결합데이터의 안전성 확보를 위해 데이터 결합시 결합데이터의 재식별 금지의무 부과 등 개인정보 재식별 방지를 위한 기술적·관리적 안전조치를 시행하여 책임성 확보에 중점을 두고 있다. 그리고 안전한 결합 환경 조성을 위해 기업 자율 데이터 결합지원, 데이터 이동권 신설과 마이데이터 정책 등 데이터결합을 지원하는 우호적인 환경조성을 위한 정책을 시행하고 있다.

구체적으로, <표 3>과 같이, 미국은 비식별화 수준을 판단하는데 HIPPA(건강보험책임법, Health Insurance Portability and Accountability Act)의 전문가 결정(Expert determination)과 18개 식별자 삭제(Safe Harbor)를 통해 프라이버시 침해 정도와 식별위험성을 합리적으로 제거하여 적정한 수준에서 데이터 활용을 위해 재식별 위험을 관리하고 있다. 또한 데이터를 결합하고 가공·판매하여 매출 1조원에 달하는 Axiom과 같은 다수의 데이터브로커(Data broker)가 결합데이터 가치를 산정하고 활발히 거래를 증가하고 있다(이대회, 2018). 데이터 개방과 Smart Closure 정책을 통해 정보주체의 자유의사에 의한 의료데이터(Blue button)와 환경데이터(Green button)의 이동이라는 마이데이터 정책도 지원하고 있다.

EU GDPR은 2016년에 여전히 개인정보에 해당하여 보호규제를 받지만, 추가정보(Additional information, 결합키)의 안전조치를 통해 개인을 알 수 없게 한 가명정보 개념을 도입하였다. 즉 추가정보의 결합이 없이는 개인을 식별할 수 없는 가명정보를 통해 '수집목적과 양립가능성' 또는 '기록보존·과학적 연구·통계목적'으로 정보주체의 동의 없이 가명정보를 활용할 수 있다. 또한 '신뢰 할 수 있는 제3자'(Trusted Third Party)를 통해 공공과 민간의

Table 3. Comparison of data combination by Country

	U.S.	EU	Japan	Korea
Legal system amendment of Data combination	<ul style="list-style-type: none"> • De-identification information removed a relevance with data subject(NIST, 2015) 	<ul style="list-style-type: none"> • Define 'Information processed that an individual can not be identified' as Anonymous Information(GDPR, 2016) 	<ul style="list-style-type: none"> • 'Anonymously Processed Information' that an individual can not be restored(Personal Information Protection Act, 2015) 	<ul style="list-style-type: none"> • Data utilization after Pseudonymization of personal information without consent by data subject
	<ul style="list-style-type: none"> • HIPPA Expert Determination, • Safe Harbor(remove 18 Identifier) 	<ul style="list-style-type: none"> • Self-check on intentional Data Re-identification, • k-Anonymity is optional 	<ul style="list-style-type: none"> • Data combination permit by Public institutions through Trusted Third Party(List Broker) 	<ul style="list-style-type: none"> • Regulation by Presidential decree
Responsibility of Data Combination	<ul style="list-style-type: none"> • Introduction of Data Re-identification Possibility Measure Test, • Re-identification Prohibition contract to Data Provider etc. 	<ul style="list-style-type: none"> • Public open restriction of De-identification information, • Information Accessor identity check, minimum verification on Re-identification, • Data destruction in case of Re-identification 	<ul style="list-style-type: none"> • Irrecoverable of personal data, Information leakage Prevention, • Announcement when providing to a third party, • Prohibition of re-identification 	<ul style="list-style-type: none"> • Keep separately 'Additional Information', • Prohibited obligation of Re-identification of Pseudonymous information, etc.
Policy to create an environment friendly for data combination	<ul style="list-style-type: none"> • Data combination by corporate autonomy • Big data R&D(2016) • AI R&D strategy(2018) • Open data(233,000 data set/year) • Smart Closure(Blue button, Green button) 	<ul style="list-style-type: none"> • Data combination Spread in EU region response to U.S. big IT tech. • Data Portability right of GDPR • U.K. ADRN(Trusted third Combining model) 	<ul style="list-style-type: none"> • National Data Combination Policy • AI comprehensive strategy plan(2019) 	<ul style="list-style-type: none"> • Competition with IT Tech • Combination support to Finance data • Open data(25,000 data set/year) • Mydata, AI/Data Plan
	<ul style="list-style-type: none"> • Free Data combination in Private sector(Data Broker) 	<ul style="list-style-type: none"> • Application of k-anonymity etc 	<ul style="list-style-type: none"> • Data combination by Public or Private institutions under supervision of Personal Information Protection Commission (Name book sale company) 	<ul style="list-style-type: none"> • Special institutions designated by Government

데이터결합을 허용하고 있고, 민간의 데이터거래는 리스트 브로커(List broker)가 데이터 결합의 가치 산정 등을 담당하고 있다(박경신, 2018; 한국인터넷진흥원, 2016).

일본은 2015년 개인정보보호법을 개정하여 복원 불가능한 가명처리에 해당하는 '익명가공정보'를 도입하여 개인정보보호위원회의 감독 하에 안전한 이용을 관리하고 있다. 가명정보와 비슷한 익명가공정보는

데이터의 활용을 위해 정보주체의 동의 없이도 개인 정보를 제공하는 등의 활용이 가능하지만, 익명가공 정보 공표의무, 정보누설방지 등 안전조치의무를 부과하고 있다. 가명정보와의 차이점은 가명정보가 여전히 개인정보에 해당하지만, 익명가공정보는 개인정보에 해당하지 않는다. 또한 '명부판매상'이 민간의 데이터 가치산정과 거래를 담당하고 있다.

한편 우리나라는 관계부처 합동으로 2016년에 ‘개인정보 비식별조치 가이드라인’을 발간하였고, 비식별 조치된 정보는 개인정보가 아닌 것으로 추정하여 데이터 분석 등에 활용할 수 있도록 하였으나, 가이드라인의 법적근거 미흡, 비식별정보를 개인정보가 아닌 것으로 추정하는 법적지위의 혼란, 비식별조치를 하더라도 개인정보라는 반증이 나올 경우 법령위반으로 처벌받을 수 있어 기업은 실효성이 낮다고 지적한다(고학수 외, 2017; 조성은, 2017). 최근에는 판매 목적의 데이터결합 금지 등 시민단체의 형사고발(2017.11)로 데이터 결합이 중단 된 상태이다. 최근 통과된 개인정보보호법 개정안에 의하면, 가명정보를 개인정보에 해당하는 비식별정보로 규정하여 과학적 연구 등의 목적으로 안전조치 후 활용하도록 하고, 구체적인 데이터 결합의 방법과 절차, 국가에 의한 결합기관의 지정, 결합데이터의 재식별 금지 등 의무 사항 등은 대통령령에 규정하도록 하고 있다. 또한 연간 25,000개의 데이터셋의 공개와 신용정보의 마 이데이터 정책도 추진하고 있다.

결론적으로, 국가간 비교정책을 통해 데이터 결합이 활성화 되는 관리적 측면의 조건을 살펴볼 수 있다. 비교 국가에서 공통적으로 데이터 결합을 지원하기 위해 비식별정보, 가명정보, 익명가공정보 등 개인정보의 개념을 명확히 하고 비식별조치와 함께 적정성 평가를 위한 제도적 사항을 법령에 규정하여 정비하고 있다. 또한 결합데이터의 오남용과 재식별위험을 방지하는 책임성 강화방안을 마련하는 한편, 마이데이터 정책, 데이터 개방, 데이터 거래소 설립 등 데이터 결합에 우호적인 환경을 조성하는 결합정책을 공통적으로 시행하고 있다.

따라서 본 연구는 데이터결합과 관련한 선행연구를 통해 기본 frame과 하위요소를 도출하였다(표 2 참조), 이에 더하여 국가간 비교정책 결과를 통해 3개의 관리적 요소를 추가로 도출한 후(표 3 참조) 이를 종합적으로 고려하여 다음의 <표 4>와 같이 총 13개 데이터 결합요인 후보군을 도출하였다.

Table 4. Candidate Factors of Data Combination

Variables	Component Factors of Data Combination		Source
	Factors	Attributes	
Technological Approach	De-identification of Data Combination	Technical proficiency that eliminates the identification of personal information before heterogeneous data combining between data and the Risk of Re-identification of combined data	Woo, S. (2018) Yang, H. (2015)
	Standardization of Combined Data	When combining data, the degree to which data types, formats, definitions, and components are processed into standardized data to be easily combined.	Yoon, S. & Oh, J. (2018), Kang, J. (2017)
	Technology Value of Data Combination	When combining data, the benefits of combining technology minus technical costs such as non-identifying measures and data processing.	Chan et al. (2013), Oliveira et al. (2014)
	Acceptance Cost of Data Combination	The cost of combining data with other data, including de-identification, data processing, and combining fees.	Lian et al. (2014) Lee, S. (2016)
	Data Combination Experience	Whether a Institution has previously performed data combination or not.	Lee, S. (2016)
Managerial Approach	Type of Data Combination Institute (Data Combination Governance)	Organizational type performing Data Combination activity such as de-identification and adequacy assessment when combining data (free type, intermediary type, designated type)	Lee, E. et al. (2017)
	Innovation Necessity of Data Combination	When data combining, The degree to which market development, product development, survey and research are actively utilized	Kim, J. & Song, T. (2014)

Manag erial Appro ach	Top management Support of Data Combination	The degree to which top management, executive team of the organization support the organization, personnel, budget, administrative management.	Chan et al. (2013) Lee, S. (2016)
	Receptivity of Data Combination	The degree to which the supplier and the consumer of data combination are accommodated by easily acquiring the information subject's favor and data quality management when combining data	Yoon, S. & Oh, J. (2018)
	Competitor Pressure of Data Combination	The pressure and intensity due to multiple competitors in the market to try to combine data	Oliveira et al. (2014), Lee, S. (2016)
	Legal amendment of Data Combination	Institutional factors such as De-identification guidelines, Privacy Acts, and big data promotion laws that support data combination.	Woo, S. (2018)
	Responsibility of Data Combination	Strength of punishment for safety management in case of violation of Data Combination such as encryption of combined data, designation of security officer of combined data, and approval of export of combined data	Kim, Y. (2019)
	Policy to create an environment friendly for data combination	Government legislation, promotion policy, pilot projects to promote the acceptance of big data when combining Personal Data	Lian et al. (2014)

IV. 연구설계 및 분석결과

4.1 델파이 조사 및 자료수집

본 연구는 빅데이터 활용에 영향을 미치는 요인으로서 개인정보 규제수준요인과 데이터 결합요인을 도출하기 위해 델파이 조사를 실시하였다. 일반적으로 델파이 기법은 전문가의 식견과 경험을 바탕으로 특정한 이슈에 대하여 원인과 문제점 그리고 미래예측을 통한 대안을 마련하는 사회과학의 한 방법론이다(이종성, 2001). 방법론상으로는 응답자의 익명성 유지(Anonymity), 설문절차의 반복(literation with controlled feedback)과 통제된 환류, 통계적 집단반응 절차(statistical group response)를 통해 의사소통을 구조화하여, 각 패널에게 이전 라운드 조사결과를 참조하여 자신의 의견을 보완해 나감으로써 궁극적으로 집단적 의사결정의 통합을 이끌어내는 점에서 차별성을 갖는다(이종성, 2001; 이경수 외, 2016). 델파이 기법에서 가장 중요한 것은 전문가의 선정과 라운드 횟수이다

이번 델파이 조사기간은 2019년 8월에 1라운드(1차 조사), 9월에 2라운드, 그리고 11월에 3라운드로 진행하였다. 1라운드 설문조사는 온라인 조사로서 주관식으로 설문하였다. 2라운드와 3라운드 조사에서는 1라운드, 2라운드에서 제시된 결과(평균, 순위)와 함

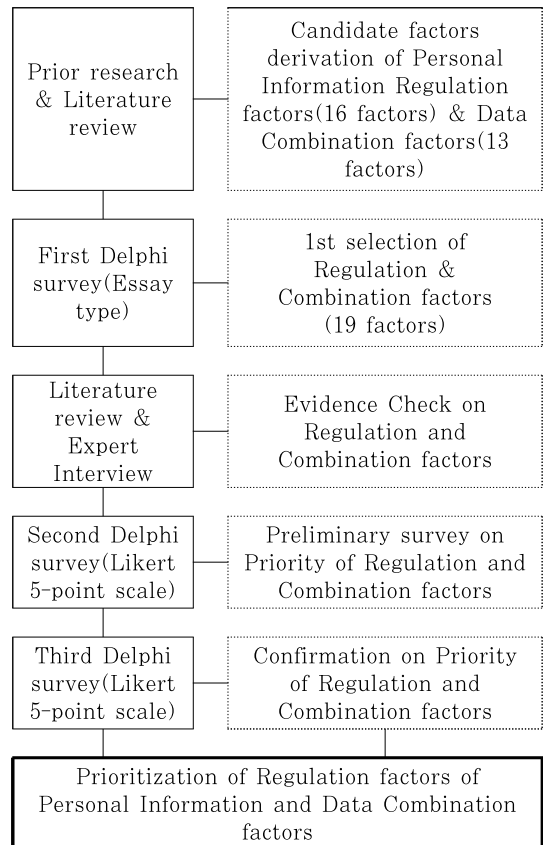


Fig. 1. Research procedure

께 5점 척도로 설문하였다.

델파이 설문내용으로서 1라운드는 개방형 질문으로서 ① 개인정보 규제수준 구성요인 ②데이터 결합 구성요인 ③ 개인정보 규제수준이 데이터 결합과 빅데이터 활성화에 미치는 영향을 설문하였다. 이어서 2라운드와 3라운드의 설문내용은 1라운드 설문결과(평균과 순위)를 참고하여 ① 개인정보 규제수준 구성요인 ② 데이터 결합 구성요인을 5점척도(1: 전혀 아니다, 2: 아니다, 3: 보통이다, 4: 그렇다, 5: 매우 그렇다)로 설문하였다.

수집된 설문의 분석은 SPSS 22.0을 사용하여 응답자 특성에 대한 빈도분석, 설문문항에 대한 평균 및 표준편차 등 기초통계조사를 실시하였다. 1차 델파이 조사는 선행연구에서 도출한 29개 후보군을 대상으로 전문가의 의견을 개방형 설문(Essay-Type)으로 조사하여 19개의 요인으로 확정하였다. 이어서 2차 델파이 조사에서는 1차 조사에서 도출된 문항에 대하여 Likert 5점척도 설문(Choice-Type)으로 선호도를 조사하여 평균과 표준편차를 구하였다. 3차 델파이 조사는 2차 조사결과를 첨부하여 전문가에게 제시하고 Likert 5점척도로 최종의견을 수렴하여 19개 요인에 대한 평균과 표준편차, 중위값 CVR, 우선순위 등을 제시하였다.

4.2 델파이 대상 전문가의 일반적 특성

델파이 조사대상은 공무원 집단, 공공기관, 교수 등 전문가 집단, 기업대표, 변호사 등 민간집단 등 3

Table 5. Delphi survey's respondent

Group	Members
Expert Group (10 persons)	Professor(Information Security, Public administration, Computer science etc.) Korea Internet Security Agency, National Information Society Agency
Government Official Group (12 persons)	Personal Information Protection Commission, Ministry of the Interior and Safety, Korea Communications Commission, Financial Services Commission etc.
Private enterprise Group (9 persons)	Corporate executive, Law Firm Kim & Chang, Lee & Ko, Asan medical center, SK telecom etc.

개 집단의 구성원으로 선정하였다. 구체적으로 개인정보보호와 빅데이터 전문가로서 교수, 시민단체 임원, 데이터 기업 임원, 변호사, 한국인터넷진흥원 등 공공기관, 행정안전부·개인정보보호위원회·방송통신위원회·금융위원회의 공무원 등 총 31명으로 구성하였다.

4.3 1차 델파이 조사결과

개방형 1차 조사에서 개인정보 규제요인에 대하여 다양한 전문가의 의견이 제시되었다. 개인정보보호법령 위반시 징역, 벌금 등 형사별로 규정되어 있어 개인정보 유출사고시 기업은 CEO 대신 보안책임자가 문책을 당한 사례가 많고, 개인정보보호법, 정보통신망법, 신용정보법 등 개인정보보호 법률이 분야별로 나뉘어져 있어서 기업입장에서 어느 법률의 적용을 받는지 혼란스럽고 중복되어 벌칙을 받는 경우의 제약점도 제시되었다. 기업이 처음 개인정보 수집시 목적과 다른 이유로 재차 개인정보 동의를 받아야 하는 경우가 많아 위탁이나 새로운 업무 시에 부담이 되고, 일부 게임업체가 유럽에 진출하고자 하는 경우에 최근 EU GDPR에 매출액의 4%에 해당하는 막대한 과징금 규정이 있어 시장에 진출하려는 의사결정을 미루고 있는 문제도 제기되었다.

또한 개인정보 개념이 모호하고 광범위하게 개인정보로 규정하고 있어 우리나라도 외국처럼 가명정보나 비식별정보 개념을 시급히 도입해야 한다는 의견도 있었다. 미국형 개인정보 규제와 유럽형 개인정보 규제를 절충해서 데이터 산업을 활성화해야 한다는 주장도 있었다. 한국은 주민등록번호로 개인정보 유출 사고가 끊이지 않으므로 빅데이터 활용을 근거로 대기업 위주 데이터 거래를 허용해서는 안 된다는 의견도 있었다. 미국은 네거티브 규제방식, Opt-out(사후 동의배제)이므로 데이터브로커 시장이 활성화 되므로 벤치마킹이 필요하다는 의견도 제기되었다. 또한 일본이 최근 EU GDPR 적정성 평가를 통과하였고 익명가공정보를 도입하여 빅데이터 육성에 나서고 있는 점은 시사하는 바가 크다고 제언하였다. 이에 본 연구는 선행연구 및 문헌조사를 통해 구성한 행위의 불확실성과 의무준수부담을 기준으로 전문가 설문을 반영하여 다음 <표 6>과 같이 1차 델파이 분석결과로 11개의 개인정보 규제요인을 추출하였다.

한편 데이터 결합요인의 1차 개방형 조사의견으로는 정부의 개인정보 비식별조치 가이드라인의 법적근

거가 행정규칙에 불과하므로 빠른 시일 내에 데이터 결합에 대한 근거규정을 개정해야 한다는 의견, 데이터 결합 전문기관으로 한국인터넷진흥원(KISA) 등 공공기관에만 한정하지 말고 문호를 넓혀서 민간 협회와 데이터 기업에게도 중개기능을 부여하자는 제안, 영국의 행정데이터네트워크(ADRN) 사례와 같이 데이터결합의 수요자, 공급자, 정부기관 어느 누구도 결합데이터의 결합키 값을 알지 못하도록 결합 거

버너스를 만들어야 한다는 의견이 있었다.

한편, 데이터 결합의 필수요건으로 데이터 결합기관 유형이 국가의 배타적으로 결합권능을 부여하는 지정형이 아닌 중개형이나 자유형에 가까울수록 데이터 결합이 증가할 것이라는 의견도 제시되었다. 또한 데이터 결합에 대한 법령위반시 형사벌을 너무 무겁게 하면 데이터 결합이 위축된다는 의견, 데이터 결합의 수요자와 공급자의 매칭을 통해 데이터 거래소

Table 6. Analysis result of 1st delphi survey

Personal Information Regulation Factors(11)		Data Combination Factors(8)	
Component Factors	Detail contents	Component Factors	Detail contents
Integration of Personal Information Supervisory Authority	Jurisdiction of Supervisory Authority, Task Convenience, Division of Supervisory Authority	De-identification of Data Combination	Re-identification Risk, Scope of Non-identification, Degree of Use contribution
Clarity of Personal Information Definition	Personal Information or not, Scope Clarity of Personal Information, Definition specification of Pseudonymous Information & Anonymous Information etc.	Technical Value of Data combination	De-identification technique, Privacy Model, Technology development
Introduction of Pseudonymous Information etc.	Compatibility with Data Collecting Purpose, Expression of Utilization Purpose such as Archiving in the public interest, Statistical purpose etc.	Data combination Experience	Combination or not, Combination intention, Introduction of Combination system
Consistency among Personal Information Acts	Jurisdiction of Act, Overlap of legal obligation, Overlap of punishment	Type of Data Combination institute	Free type, Intermediary type, Designated type
Clarity of Data combination Regulation	De-identification regulation or not, Data Combination Procedure & Effect	Standardization of Combined data	Consistency of Data type, Definition, Component
Evidence Clarity of Personal Information De-identification	De-identification regulation, Privacy Model, De-identification technique	Responsibility of Data combination	Encryption of Combining Data, Safety measures, Punishment intensity
Negative legislation of Permission principle	Legislation elasticity, Legislation method	Legal Amendment of data combination	De-identification Guideline, Degree of Statute
Ease of Personal Information Consent	Severity of Prior Consent, Consent Safety, Consent Convenience	Policy to create an environment friendly for data combination	Guideline, Promotion policy, Pilot project.
Accordance of Consent Requirements between Personal Information Collection & Use and Offering	Collection & Use condition, Offering condition to Third Party		
Proper punishment Intensity in case of violation of law	Administrative punishment intensity, Criminal punishment intensity, Civil Responsibility		
Proper Penalty Level when comparing EU GDPR	Imposition requirement of Penalty, Penalty Scale		

와 결합플랫폼을 만들어야 한다는 의견 등이 제안되었다. 또한 정부 가이드라인에 의한 비식별조치는 비용이 너무 많이 들고 최종책임을 기업에게 지우고 있어 일정조건하에 기업의 면책조항을 부여해야 데이터 결합이 활발해질 것이라는 의견, 개인정보 정의, 동의

요건 등 엄격한 사전규제는 데이터 결합의 규제와 연계되어 있으므로 사후규제(opt-out)를 하는 방향으로 법률 개정이 필요하다는 주장도 있다. 이어서 우리나라 네티즌 특공대는 다양한 출처로 부터 개인의 데이터를 조합해서 특정 개인을 식별해 내므로 결합

Table 7. Analysis result of 2nd and 3rd Delphi survey

Variables	Regulatory Level Component Factors	Round	Mean	Standard Deviation	Median	Quartile Average		CVR	t-value	p-value	Priority
						Q1	Q3				
Personal Information Regulation Factors (11)	Introduction of Pseudonymous Information etc.	2 nd	4.10	.615	4.00	4.00	4.00	0.935	-1.511	.142	1
		3 rd	4.39	.558	4.00	4.00	5.00				
	Evidence Clarity of Personal Information De-identification	2 nd	4.13	.819	4.00	4.00	5.00	0.741	-.724	.475	2
		3 rd	4.32	.909	5.00	4.00	5.00				
	Clarity of Data Combination Regulation	2 nd	4.07	.907	4.00	4.00	5.00	0.806	-.796	.433	3
		3 rd	4.26	.815	4.00	4.00	5.00				
	Clarity of Personal Information Definition	2 nd	4.03	.999	4.00	4.00	5.00	0.612	-.432	.669	4
		3 rd	4.13	.718	4.00	4.00	5.00				
	Ease of Personal Information Consent	2 nd	4.00	1.05	4.00	3.75	5.00	0.677	-.232	.818	5
		3 rd	4.06	.854	4.00	4.00	5.00				
	Integration of Personal Information Supervisory Authority	2 nd	3.80	.997	4.00	3.00	4.25	0.612	-.783	.440	6
3 rd		4.03	.875	4.00	4.00	5.00					
Consistency among Personal Information Protection Acts	2 nd	4.03	.615	4.00	4.00	4.00	0.612	.421	.677	7	
	3 rd	3.97	.706	4.00	4.00	4.00					
Accordance of Consent Requirements between Personal Information Collection & Use and Offering.	2 nd	3.87	.900	4.00	3.00	4.25	0.354	.000	1.000	8	
	3 rd	3.87	.922	4.00	3.00	5.00					
Adequacy Punishment Intensity in case of violation of Law	2 nd	3.90	.995	4.00	3.00	5.00	0.354	-.141	.889	8	
	3 rd	3.87	.991	4.00	3.00	5.00					
Proper Penalty Level when comparing EU GDPR	2 nd	4.07	.740	4.00	4.00	5.00	0.419	.668	.509	8	
	3 rd	3.87	1.02	4.00	3.00	5.00					
Negative legislation of Permission Principle	2 nd	4.10	.995	4.00	4.00	5.00	0.290	1.188	.245	-	
	3 rd	3.84	.898	4.00	3.00	5.00					
Data Combination Factors (8)	De-identification of Data Combination	2 nd	4.07	.640	4.00	4.00	4.25	0.548	.360	.509	1
		3 rd	4.03	.912	4.00	4.00	5.00				
	Standardization of Combined Data	2 nd	4.13	.900	4.00	4.00	5.00	0.419	.538	.595	1
		3 rd	4.03	1.11	4.00	3.00	5.00				
	Responsibility of Data Combination	2 nd	4.07	.868	4.00	4.00	5.00	0.483	.258	.798	3
		3 rd	3.97	.912	4.00	3.00	5.00				
	Type of Data Combination Institute	2 nd	4.07	.691	4.00	4.00	5.00	0.548	.924	.363	4
		3 rd	3.87	.885	4.00	4.00	4.00				
	Data Combination Experience	2 nd	4.00	.830	4.00	3.00	5.00	0.354	.634	.531	5
		3 rd	3.84	1.03	4.00	3.00	5.00				
	Technological Value of Data combination	2 nd	3.87	.571	4.00	3.75	4.00	0.419	.328	.745	6
3 rd		3.77	.805	4.00	3.00	4.00					
Legal Amendment of data combination	2 nd	4.13	.819	4.00	4.00	5.00	0.225	1.578	.125	-	
	3 rd	3.74	.965	4.00	3.00	4.00					
Policy to create an environment friendly for data combination	2 nd	4.00	.743	4.00	4.00	4.25	0.225	1.246	.223	-	
	3 rd	3.74	1.09	4.00	3.00	5.00					

데이터를 통해 개인을 재식별하는 경우에 엄벌에 처하는 것이 필요하다는 제안도 있었다. 이에 본 연구는 데이터 결합의 기술적·관리적 접근방법과 국가별 비교연구를 기반으로 1차 전문가 설문을 반영하여 다음 <표 6>과 같이 8개의 데이터 결합요인의 후보군을 도출하였다.

4.4 2차 및 3차 델파이 조사결과

델파이 2차 조사에서는 <표 7>에서 보는 바와 같이, 개인정보 규제요인 11개 요인을 리커트 5점 척도 객관식 질문을 통해 평균과 우선순위를 추출하였다. 2차 조사결과, 개인정보 규제요인으로 개인정보 비식별화 근거 명확(4.13), 가명정보 등 활용제도 도입(4.1), 네거티브 원칙허용 입법양식(4.1), 데이터 결합 규정 명확성(4.07), GDPR과 비교시 적정 과징금(4.07), 개인정보 법률간 정합성(4.03), 개인정보 정의 명확성(4.03), 개인정보 사전동의 용이성(4.0), 법령위반시 적정 처벌강도(3.9), 개인정보 수집이용과 제공간 동의요건 일치(3.87), 개인정보 감독기구 통합(3.8) 순으로 나타났다.

한편 데이터 결합요인에 대한 2차 조사는 TOE 프레임워크에서 사용된 요인 중 8개의 데이터 결합요인을 추출하였다. 우선순위가 높은 기술·조직·환경요인으로 결합데이터 표준화(4.13), 결합 법령개정(4.13), 결합 비식별성(4.07), 결합 책임성(4.07), 결합기관 유형(4.07), 결합 경험(4.0), 데이터 결합에 우호적인 환경조성 정책(4.0), 결합 기술가치(3.87) 순으로 조사되었다.

개인정보 규제요인에 대한 3차 조사결과에 따른 최종 우선순위는 ①가명정보 등 활용제도 도입 ②개인정보 비식별화 근거 명확성 ③데이터 결합규정 명확성 ④개인정보 정의 명확성 ⑤ 개인정보 사전동의 용이성 ⑥개인정보 감독기구 통합 ⑦개인정보 법률간 정합성 ⑧개인정보 수집이용과 제공간 동의요건 일치 ⑧법령위반시 적정 처벌강도 ⑧GDPR 비교시 적정 과징금 순으로 나타났다. 한편 데이터 결합의 3차 조사결과 최종순위는 ①결합 비식별성 ②결합데이터 표준화 ③결합 책임성 ④결합기관 유형 ⑤결합 경험 ⑥결합 기술가치 순으로 조사되었다.

V. 결론 및 시사점

본 연구는 개인정보보호 실태조사와 규제요인 관련

연구를 통해 개인정보 규제요인을 도출하는 한편 데이터 결합의 접근방법 및 국가별 데이터 결합 비교를 통해 데이터 결합요인을 도출하였다. 이를 토대로 델파이 분석을 시도하여 빅데이터 활성화에 영향을 미치는 개인정보 규제요인과 데이터 결합요인이 무엇인지 그 우선순위를 시론적으로 탐색하였다. 델파이 분석결과, 첫째, 개인정보 규제요인은 ①가명정보 등 활용제도 도입 ②개인정보 비식별화 근거 명확성 ③데이터 결합규정 명확성 ④개인정보 정의 명확성 ⑤ 개인정보 동의 용이성 ⑥개인정보 감독기구 통합 ⑦개인정보 법률간 정합성 ⑧개인정보 수집이용과 제공간 동의요건 일치, 법령위반시 적정 처벌강도, EU GDPR 비교시 적정 과징금 순으로 나타났다. 둘째, 데이터 결합요인은 한편 데이터 결합의 3차 조사결과 최종순위는 ①결합 비식별성 ①결합데이터 표준화 ③결합 책임성 ④결합기관 유형(결합 거버넌스) ⑤결합 경험 ⑥결합 기술가치 순으로 조사되었다.

이러한 연구결과는 빅데이터 활용을 위한 개인정보 규제와 데이터 결합정책 설계시 바람직한 정책방향과 어떤 정책과제에 우선하여 추진해야 하는지 시사점을 제공한다. 첫째, 개인정보 규제요인으로 높은 우선순위를 보인 가명정보의 도입 등 개인정보의 정의를 명확히 하고, 비식별화(가명처리) 및 데이터 결합에 대한 세부사항을 명확히 제시하여야 한다. 최근 개정된 개인정보보호법 개정안(제28조의2, 제28조의3, 제28조의4)에는 가명정보의 활용범위를 통계작성, 과학적 연구, 공익적 기록보존으로 한정하고 있으며, 데이터 결합의 구체적 방법과 절차, 국가의 전문기관 지정, 안전조치 등을 대통령령에 위임하고 있다. 데이터경제 활성화의 핵심이 가명정보의 활용범위와 데이터 결합에 있으므로 통계와 과학적 연구에 대한 범위와 해석을 명확히 제시하여야 한다. 또한 결합된 데이터 활용을 위한 데이터 표준화 및 결합된 데이터를 고의적으로 또는 결합과정에서 우발적으로 재식별하는 경우의 위험요인을 최소화하는 구체적인 결합절차와 결합기관 거버넌스 등 가이드라인을 제시하여야 한다.

둘째, 개인정보 감독기구의 통합은 최근 개정된 개인정보보호법에 의하면, 행안부, 방송통신위원회, 금융위원회로 나뉘어진 감독기구를 개인정보보호위원회로 일원화하여 개인정보 유출과 오남용에 대한 안전장치를 효율화하였다. 다만, 산업계에서 지적하듯이, 개인정보보호위원회가 보호의 컨트롤타워로서 보호규제 일변도로 흐르지 않도록 위원회 구성시에 데이터 활용을 대표하는 인사를 위원으로 위촉하여 보호와

활용의 조화를 이루도록 해야 한다.

셋째, 델파이 분석결과에서 높은 순위를 보인 개인정보 동의 용이성, 개인정보 수집이용과 제공간 동의 요건 일치, 법령위반시 적정 처벌강도 등은 최근 개인정보보호법 개정안에 충분히 반영되지 않았으므로 향후 법령개정이 필요하다. 먼저, 정보주체의 엄격한 사전동의를 개선할 필요가 있다. 2016년 개인정보보호 실태조사에서 사전동의시 정보주체가 동의서를 확인하고 의사표시를 하는 비율이 23.2%에 불과하다는 점(이상무, 2019)을 고려하면 포괄적 동의 또는 사후동의 배제방안의 도입을 검토해야 한다. 또한 개인정보 수집이용시 동의요건과 제3자 제공시 동의요건이 달라서 수법자의 혼선과 데이터 경제 활용에 제약요인이 되고 있다. 그러므로 개인정보보호법 제17조 제3자 제공 동의요건에 정보주체와의 계약체결이 불가피한 경우와 개인정보처리자의 정당한이익이 정보주체의 권리보다 앞서는 경우에는 제15조 개인정보 수집·이용시 동의요건과 일치하도록 법률개정이 필요하다. 한편 개인정보 법령위반 시 과도하게 형사벌 위주로 되어있는 처벌규정을 개선할 필요가 있다. 현행 개인정보보호법에 의하면 제3자에게 정보주체의 동의없이 개인정보를 잘못 제공하는 경우에 징역5년 이하이 징역에 처해질 수 있고, 기업대표나 임원은 범위반으로 징역형을 선고받을 수 있어 과도한 형사처벌이 문제가 된다(김용학, 2019). 그러므로 과도한 형사벌을 낮추고 EU GDPR의 과징금 강화정책과 같은 경제벌로 전환하는 제도개선이 필요하다.

넷째, 데이터 결합 거버넌스와 관련하여 데이터 경제의 시너지 창출을 위해 데이터 결합 중개기관의 기능을 데이터 개방기능과 데이터 거래기능과 연계하여야 한다. 개정된 개인정보보호법에는 국가지정 전문기관이 이종기업간 데이터 결합을 수행하도록 하고 있다. 데이터 결합의 부가가치 창출을 지원하기 위해 국가지정 전문기관에 데이터 개방기능(예: 공공빅데이터 센터)과 데이터 거래기능(예: 빅데이터 거래소 플랫폼, 마이데이터)을 연계하여, 데이터 개방에서 데이터 결합, 데이터 거래까지 윈스톱으로 Data startup을 지원하는 활용체계를 만들어야 한다.

그동안 개인정보의 규제개혁은 빅데이터 활성화를 논의하는데 핵심적인 요인이었음에도 불구하고 구체적인 논의가 충분하지 않았다. 때문에 개인정보의 규제개혁은 현재까지도 모호하고 추상적인 수준에 머물러 왔다. 본 연구는 개인정보 규제수준과 데이터 결합에 대한 구성요인이 무엇인지 구체적으로 제시했

다는 점에서 학술적·실무적 의미가 있다. 또한 정책적으로는 '안전한 개인정보 이용'을 위한 개인정보 규제요인이 무엇인지를 밝히고 있어 제도개선의 방향과 내용을 제공하고 있다. 다만, 연구의 한계점으로는 선행연구가 부족한 상태에서 개인정보 규제요인과 데이터 결합요인의 개념화를 시도하는데 그치고, 각각의 요인이 어떠한 경로를 거쳐 빅데이터 활용에 영향을 미치는지 명확하고 풍부한 인과관계 분석에 이르지 못하였다.

그러므로 향후 연구방향은 먼저, 개인정보 규제요인과 데이터 결합요인에 대한 타당성 조사와 국내외 2차적 통계데이터를 통한 정량적 분석 등을 통해 '규제강도 결정요인' 등 표준화된 방법론을 활용하여 객관화된 요인분석이 필요하다. 또한 본 연구결과를 토대로 개인정보 규제수준과 데이터 결합이 빅데이터 활용에 미치는 영향에 대한 가설을 설정하고 변수의 인과관계를 규명하는 실증연구가 필요하다.

References

- [1] Jung-muk Kang, "An Analysis on the Factors Affecting of Big data Utilization in Local Government: Focused on the Public Servants Perception in Seoul," *Journal of Governance Studies*, 12(1), pp. 161-197, Mar. 2017.
- [2] Personal Information Protection Commission and Ministry of the Interior and Safety, "2018 Survey of Personal Information Protection in Korea." pp. 1-226, Feb. 2018.
- [3] Hak-soo Ko and Dong-jin Lee and Sun-goo Lee and Eun-su Kim and Jong-gu Jeong, "De-identifying Personal Information," Pakyoungsa, Seoul, pp. 1-282, Jul. 2017.
- [4] Tae-eon Koo, "Regulatory rationalization plan for innovative industry development," *Korean Federation of Science and Technology Societies*, pp. 11-33, Sep. 2017.
- [5] Young-jun Kwon, "Thought on the Self-Determination Right to Personal

- Information and the Consent Regime.” 2015 Naver Privacy White Paper. pp. 78-148, Feb. 2016.
- [6] Joint Ministries, “Personal Information De-Identification Guideline,” pp. 1-76, Jun. 2016.
- [7] Kyung-Hwan Kim, “Comparison of Personal Information Protection Act in Korea, EU, and Japan in terms of Regulation,” 2017 Naver Privacy White Paper, pp. 119-158, Dec. 2017.
- [8] Sun-kyung Kim, “A Study of the Personalization Service and Privacy Paradox in the Big Data Era-Focus on the Socio-technical Perspective,” Journal of the Korean Cadastre Information Association, 16(2), pp. 193-207, Aug. 2014.
- [9] Soo-yeon Kim, “Regulatory Improvement Review on the Personal Information Protection for Activating the Big data Industry,” KERI Brief, pp. 15-28, Dec. 2015.
- [10] Young-sin Kim and Young-joo Song, “Comparison and Implications of Domestic ICT Competitiveness: Focusing on ICT Regulation Level,” Policy Research, Korea Economic Research Institute, pp. 1-57, Dec. 2016.
- [11] Yong-hak Kim, “A Study on the Improvement of Punishment System for Personal Information Violation,” Doctoral dissertation, Soongsil University, pp. 1-159, Aug. 2019
- [12] Il-hwan Kim, “A Study on the Role and Function of Personal Information Protection Commission,” Study on the American Constitution, 28(2), pp. 219-273, Aug. 2017.
- [13] Jung-wook Kim and Kyung-sik Ki, “Economic Analysis on Regulatory Reform,” Journal of Regulation Studies, 25(2), pp. 45-79, Dec. 2016.
- [14] Jin-young Kim, “Review the Regulation and the Improvement in the Big Data: Focusing on the Policy Direction and Legislation of the New Government,” IT & Law Review, 15, pp. 157-191, Aug. 2017.
- [15] Tae-yoon Kim and Soo-a Lee, “Regulatory Stringency: with the Application of Report Regulations in Korea,” Journal of Regulation Studies, 21(2), pp. 147-185, Dec. 2012.
- [16] Korean Chamber of Commerce, “Institutional Improvement Survey on Personal Information Protection,” pp. 1-33, Nov. 2018.
- [17] Kyung-sin Park, “Big data Possibility in Personal Information Protection Legal system,” (ed.) Personal Information in the 4th Industrial Revolution Era, Sechang publisher, Seoul, pp. 196-229, Dec. 2018.
- [18] Noh-hyung Park, “Improvement Review of Korea’s Personal Information Protection Legal system according to the Analysis of Personal Information Protection Act in major countries on Big Data,” 2016 Naver Privacy White Paper, pp. 1-44, Dec. 2016.
- [19] Noh-hyung Park and Myung-hyun Jung, “Improvement of the Personal Information Protection Act for Activating Big Data Analytics: Focusing on Comparative Analysis with the EU GDPR,” Korea Law Review, 85, pp.1-39, Jun. 2017.
- [20] Young-jin Shin, “A Study on Precautionary Regulatory Status Analysis and Improvement Plan for Personal Information Protection in Korea,” Autumn Conference Booklet in the Korean Association for Policy Studies, pp. 17-41, Sep. 2018.
- [21] A. Weigend, “Data for the people: how to make our post-privacy economy work for You,” Trans. Ji-young Hong,

- Four Season, Paju, pp. 1-440, Nov. 2018.
- [22] Hyun-cheol Yang, "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, pp. 1-136, Feb. 2015.
- [23] Jung-hyun Oh and Mi-ran Sun, "Survey Analysis on Limitation Cases of Big data Utilization due to Personal Information Legal System." *Big Data*, Issue 4, pp. 1-18, Jun. 2015.
- [24] Soon-kyu Woo, "A Study on the Factors Affecting Personal Information De-identification Based on Big Data in the Financial Industry," Doctoral dissertation, Soongsil University, pp. 1-152, Aug. 2018.
- [25] Kwang-suk Yoon, "A Study on Administrative Service Innovation using Information Technology in 4th Industrial Revolution Era," Research Paper of 'Korea Institute of Public Administration', 2018-16, pp. 1-428, Dec. 2018.
- [26] Soo-young Yoon and Jae-in Oh, "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), pp. 67-94, Nov. 2018.
- [27] Hye-seon Yoon, "A Search for a new Regulatory Paradigm for Big data" *Journal of Law and Economic Regulation*, 11(1), pp. 71-94, May. 2018.
- [28] Kyung-soo Lee and Tae-hyung Kim and Tae-hoon Kim and Sang-hyuk Park, "Analysis of Risk Factor for Sinkhole Formation by using Delphi," *Journal of the Korea Contents Association*, 16(4), pp. 65-75, Apr. 2016.
- [29] Dae-hee Lee, "A Study on Pseudonymized and De-identified Information for the Protection and Free Movement of Personal Information," *Journal of Korea Information Law*, 21(3), pp. 217-251, Dec. 2018.
- [30] Min-hwa Lee, "The 4th Industrial Revolution in Korea," *Creative Economy Association*, Seoul, pp. 1-446, Jun. 2017.
- [31] Sang-moo Lee, "Personal Information Utilization and Role of Government," (ed.) *Regulatory Reform Task and Case Study for promoting New Industries*. Jinhan M&B, Seoul, pp. 199-254, Jun. 2019.
- [32] Sang-jik Lee, "Problem between the 4th Industrial Revolution and Personal Information Protection: Focusing on EU GDPR," (ed.) *Personal Information in the 4th Industrial Revolution Era*, Sechang publisher, Seoul, pp. 22-55, Dec. 2018.
- [33] Sun-woo Lee, "Research on Determinants for Big data System Adoption in Organization," Doctoral dissertation, Sungkyunkwan University, pp. 1-131, Aug. 2016.
- [34] Seong-yeop Lee, "Suggestion of Improvement Plan for advancement of Personal Information Protection Act." *Conference Booklet of 50th Anniversary of E-Government*, Oct. 2017.
- [35] Eun-woo Lee, "Improvement Task of Personal Information Protection Act in 20th National Assembly," *Booklet of Big data Utilization and Personal Information Protection*. Future Creation Science Broadcasting & Communications Commission in National Assembly, Jul. 2017.
- [36] Eun-woo Lee and Byung-il Oh and

- Yeo-kyung Jang and Kyu-man Hwang, "A Study on Introduction of Data Linkage & Combination Support System," Personal Information Protection Commission, pp. 1-460, Dec. 2017.
- [37] Jong-sung Lee, "Delphi Method," Education Science company, Seoul, pp. 1-138, Feb. 2001.
- [38] Hye-young Lee, "A Study on Legislation of Social Regulation: the Perspective of Regulatory Paradox," Journal of Parliamentary Research, 13(1), pp. 5-69, Jun. 2019.
- [39] Seong-En Cho, "Improvement Strategy of Information Usability under Personal Information Protection Legal System," Premium Report of Korea Information Society Development Institute, pp. 1-28, Oct. 2017.
- [40] Kyung-jin Choi, "Artificial Intelligence, Data, and Law," (ed.) Data Economy, Chapter 11, Hans media, Seoul, pp. 245-266, Jul. 2017.
- [41] Jin-wook Choi, "Effects of Regulation on the Competitiveness of a Nation in OECD Countries," Journal of Regulation Studies, 15(1), pp. 3-24, Jun. 2006.
- [42] Hyun-woo Choi and Yoon-jung Choi and Sung-ho Lee, "A Study on Improvement of Personal Information Protection System," Booklet of Sustainable Growth Initiative, Nov. 2018.
- [43] Seong-rak Choi, "Regulation Encyclopedia in Korea," Paper Road, Seoul, pp. 1-248, Dec. 2018.
- [44] Korea Data Agency, "2018 Data Industry White Papers," pp. 1-79, Sep. 2018.
- [45] Korea Internet & Security Agency, "A Study on Global Personal Information Utilization Case and Legal System Trends in the New ICT Environment," KISA Research Paper, pp. 1-213, Oct. 2016.
- [46] S. Chan and L. Chong, "Determinants of mobile supply chain management system diffusion: a structural equation analysis of manufacturing firms," International Journal of production Research, vol. 51, no. 2, pp. 1196-1213, Feb. 2013.
- [47] IDC, "Worldwide Semiannual Big Data and Analytics Spending Guide," Feb. 2017.
- [48] IMD, "IMD World Digital Competitiveness Ranking 2018," pp. 1-179, Jun. 2018.
- [49] W. Lian and C. Yen and T. Wang, "An exploratory study to understand the critical factors affecting the decision to adopt cloud computing in Taiwan hospital," International Journal of Information Management, vol. 34, pp. 28-36, Jan. 2014.
- [50] Luc Rocher et al, "Estimating the success of re-identifications in incomplete data sets using generative models," Nature communications, pp. 1-9, Jul. 2019.
- [51] OECD, "Government at Glance 2017 edition: Open Government," pp. 1-282, May. 2017.
- [52] T. Oliveira and M. Thomas and M. Espadanal, "Assessing the determinants of cloud computing adoption : An analysis of the manufacturing and services sectors," Information and Management, vol. 51, pp. 497-510, Jul. 2014.

〈 저자 소개 〉



김 상 광(Sang-Gwang, Kim) 정회원
2019년 2월: 충남대학교 국가정책대학원 박사과정 수료
2005년 3월~현재: 행정안전부 근무
<관심분야> 정보정책, 도시행정, 정부혁신·지역혁신



김 선 경(Sun-Kyung, Kim) 정회원
2005. 9월~현재: 충남대학교 행정학부 근무
<관심분야> 정보정책, 도시행정(관리)