이학석사 학위논문

개인정보보호법상 가명처리에 관한 연구

2023년 2월

서울대학교 융합과학기술대학원

수리정보과학과 정보보호 및 디지털포렌식학전공

성 기 범

개인정보보호법상 가명처리에 관한 연구

지도교수 천 정 희

이 논문을 이학석사 학위논문으로 제출함

2022년 12월

서울대학교 융합과학기술대학원

수리정보과학과 정보보호 및 디지털포렌식학전공

성기범

성기범의 석사 학위논문을 인준함

2023년 1월

위 원 장 _____ (인)

부 위 원 장 _____ (인)

위 원 _____ (인)

논 문 초 록(국문)

개인정보보호법상 가명처리에 관한 연구

기계 학습, 인공지능 개발 등 새로운 기술의 연구, 개발, 적용이 주목받는 소위 '4차 산업혁명'의 흐름에 따라 개인정보의 적법하고, 효율적인 이용을 촉진하고자 2020. 2. 4. '데이터 3법'이 개정되었다. 그 중 개인정보보호법은 개정을 통하여 가명정보, 가명처리의 개념을 도입하고 일정한 요건에 따라 정보주체의 동의 없는 이용을 허용하게 되었다.

헌법재판소가 정보주체의 헌법상 권리로 개인정보자기결정권을 확인하였고, 이를 보장하기 위한 입법적 조치가 개인정보보호법이라고 볼 수 있으나, 한편 정보이용 주체가 개인정보를 처리, 활용함으로써 행사하게 되는 재산권의 행사, 학문의 자유 등도 개인정보보호법에 의하여 실현될 것이어서 양자의 조화로운 보장 방안을 모색하는 것이 본 연구의 목적이다. 결국 개인정보의 적정한 이용을 통하여 정보주체의 개인정보 침해에 대한 우려를 불식시키는 것이 개정 개인정보보호법의 해석, 적용, 집행의 주요한 과제로 볼 수밖에 없는 셈이다.

개정 개인정보보호법의 해석 상 정보주체의 동의 없이 개인정보를 적법하게 처리하고자 한다면 안전성의 확보에 대한 필요한 조치(법 제15조 제3항 등)를 하거나, 해당 정보가 가명처리(법 제2조 제1호의2)를 통하여 가명정보(법 제28조의2 이하)가 되어야 하는데, 어느 경우든 비식별조치로서의 가명처리, 암호화조치가 불가피하다.

가명처리의 기법 중 하나인 수리암호에 따른 암호화조치는 수학적 난제에 기한 안전성 내지 기밀성, 컴퓨터 연산 기능 향상에 따른 효율성, 디지털데이터의 형식으로 처리된 빅데이터에 대한 가공 용이성 측면에서 우수성이 있다고 판단되고, AES, 해시함수, 동형암호와 같이 데이터 형태, 처리환경 및 목적에 따라 다양한 선택이 가능하다.

그러나 가명처리, 가명정보의 개념을 야심차게 도입한 신법에서는 적정한 가명처리의 기준, 추가 정보의 개념 등 필수적인 사항을 제대로 규율하지 않거나 하위 법령에 대한 위임을 포기하였고, 법률로부터 일부 사항에 관하여 위임 받은 시행령조차 이를 그대로 고시에 재위임한 결과, 고시 차원에서 불

충분하거나 부적절한 규율이 이뤄졌다.

이에 따라 각 행정관청은 소관 개인정보에 대한 가명처리에 관한 가이드라인을 제작, 공표하였는데, 가이드라인의 법적 성격, 규율 형태에 비추어 규범력이 희박함에도 위 가이드라인들은 수범자들에게 권장, 설명이라는 외관을 취하면서 사실상 의무를 부과하고 있어 논란이 있다.

또한 가이드라인에서는 가명처리된 식별자라고 하여도 삭제하여야 한다거나, 가명정보의 이용은 가명처리 목적, 환경에 기속된다거나, 안전한 가명처리가 개발될 때까지는 가명처리를 금지한다는 등으로 법률의 내용, 입법취지에 반하는 설명을 하고 있고, 민감정보의 가명처리, 의료법 등 단행 법률과의 관계 등 입법으로 해결할 사항에 대하여 임의로 규율을 시도하고 있는 문제점도 발견된다.

한편, 신법이 개정되는 과정에서 선례가 되었던 유럽연합의 GDPR, 미 연방 법률인 HIPAA에서는 식별가능성의 판단에 대하여 법령의 단계에서 상세한 규율을 하여 국내 법제 개선에 중요한 시사점을 제공하고 있다.

GDPR은 ▲ 식별가능성을 판단하는 주체를 개인정보처리자 외의 제3자로 넓히고, ▲ 현재 사용하는 최신의 기술, 식별에 사용되는 비용, 시간을 감안하여 식별가능성을 판단하도록 하고 있으며, ▲ 이에 더하여 개인정보 처리의 성격, 범위, 내용을 종합하여야 한다고 규율하고 있고, HIPAA는 식별자의 개념을 명시적으로 제시하면서 ▲ 아예 식별자를 삭제하거나, ▲ 통계학적, 과학적 기법을 사용하는 전문가가 합리적으로 사용가능한 정보를 결합하여도 식별가능성이 매우 낮다는 의견을 개진하는 방법으로 판단할 수 있다고 규정한다.

물론 구법 당시에도 정보주체의 동의 없이 개인정보를 이용하고자 비식별조치를 한 사례 또는 개인정보 유출 사고가 발생하였을 때 개인정보로 볼 수 없다는 주장을 한 사례가 있어 식별주체, 식별가능성에 대해 하급심이 판단하였는데, 개인정보보호에 관심을 가지고 식별주체를 제3자에까지 확대한 판결도 있었으나, 암호화조치에 대한 이해가 부족하여 개인정보처리자들의 주장에 경도된 나머지 부적정한 암호화조치를 적법하다고 보거나, 정보주체와 암호문을 연결하는 매칭테이블을 주고받은 행위를 오히려 범의를 부정하는 논거로 드는 등으로 사실을 오인하고 법리를 오해한 사례도 찾을 수 있었다.

위와 같은 판결의 문제점은 구법이 비식별조치를 명확하게 규율하지 않은 가운데 가이드라인으로서 일응의 기준만 제시하여, 법원의 규범적 판단에 관

한 최소한의 기준이 입법으로 마련되지 않은 것에 가장 큰 이유가 있다고 평가할 수 있다.

본 연구를 통하여 신법이 개정, 시행된 이후에도 이러한 구법상의 문제점은 여전히 상당 부분 잔존하고 있음을 확인하였으므로, 가명처리와 암호화조치에 관하여 입법상의 개선, 가이드라인의 보완이 필요하다는 결론에 이르렀다.

우선 입법 조치로는, ① 적정한 가명처리, 암호화조치의 판단 기준, ② 가명처리 시 사용되는 추가 정보의 개념, ③ 법 제15조 제3항, 제17조 제4항에 관한 하위 법령 단위에서의 규율, ④ 민감정보에 대한 가명처리 여부에 관한 정리, ⑤ 의료법 등 단행 법률과 개인정보보호법과의 관계 설정 등이 필요하다.

시행령 이하 하위 법령에서는 ① 위 각 입법사항에 대한 구체적인 규율에 더하여, ② 추가 정보, 암호화조치 등 가명처리의 기법에 관한 개략적인 예시 등을 정할 수 있을 것이고, 가이드라인에 대하여는 ① 가이드라인의 제작, 공표에 관하여 법령에서 일정한 사항을 위임하거나, 최소한 행정청에서 이를 정할 수 있는 근거를 마련하고, ② 법령의 취지에 상충되거나 법령이 규율하여야 할 부분은 정리하며, ③ 추가 정보의 관리, 암호화 알고리즘의 장단점 및 구현 방식 등 가이드라인에서 설명이 필요한 부분을 보완하면 좋을 것이다.

한편 기술적 측면을 고려할 때, 암호화조치가 가명처리의 여러 기법 중 하나이지만, 개정법이 상정하고 있고 실제 정부와 기업이 기도하고 있는 가명처리의 형태는 대량의 디지털데이터에 대하여 빈번한 연산을 통한 처리, 즉 '빅데이터'에 대한 처리가 될 것이므로, 암호화 알고리즘에 따라 소프트웨어로 구현되기 쉬운 수리암호화조치가 안정성 및 효용성의 측면에서 중요한 위치를 차지하고 있다.

본 연구 결과, 암호화조치의 중요성에 비하여 개인정보보호위원회가 가이드라인 등에서 암호화조치에 대하여 구체적인 설명을 회피하거나 부적절, 불충분한 설명에 그쳤고, 일부 법원 또한 비식별조치의 적정성을 판단하는 과정에서 암호화조치를 하나의 요건으로 취급하는데 그쳐 적절한 판단을 그르친 사례가 확인되는데, 이는 암호화조치의 개념이나 중요성은 물론 기술적 조치에 관한 이해가 부족한 것에 이유가 있다고 보았다.

예를 들어 이미 널리 활용되고 있는 해시함수 등 일방향 암호화에 대하여

다양한 공격방법을 상정하여, 매칭테이블 작성, 보관을 엄격히 금지하고, 사전계산을 통한 공격에 대비할 수 있는 솔트와 같은 기술적 조치의 근거를 마련할 필요가 있으며, 잡음(노이즈) 추가의 경우 수리암호화에서 사용하는 수준의 보안성을 요구할 필요가 있을 것이다.

가이드라인이 삭제, 부분 삭제, 마스킹을 가명처리 기법으로 들고 있다고 하더라도 법령 단계에서 식별가능성의 판단 기준을 정립하지 않는 이상 정보이용주체가 이를 악용하여 불완전한 가명처리에 그칠 염려가 있고, 순열, 라운딩과 같이 정보이용가치를 현저히 저감하게 하는 가명처리는 정보이용주체가 선택할 가능성이 희박하므로, 위와 같이 적정한 가명처리의 판단 기준을 명문화하는 기회에 암호화 알고리즘별 장단점, 적용 가능한 정보의 형태, 분야를 범주화하는 노력도 필요하다고 생각된다.

덧붙여 개인정보의 유출 등 사고가 발생되어 정보이용주체가 더 이상 가명처리된 정보를 관리하지 못하게 되는 상황에서도 유효한 가명처리 기법을 규율할 필요가 있고, 특히 암호화조치에 대한 규범적 판단을 내림에 있어, 식별가능성 판단 기준을 강화하면서 특히 기술의 특성을 고려한 정치한 논증이 필요할 것이다.

정보주체인 사회 구성원들이 개인정보의 적정한 이용에 공감대를 가져야만, 정보이용 주체들은 빅데이터 이용을 통한 기술의 연구, 개발 및 구현에 적극적으로 나설 수 있게 된다.

본 연구에서 정리한 문제점, 개선 착안점 등 제언한 내용 등이 후속 연구와 입법 등으로 이어져, 행정, 사법기관이 균형 있게 개인정보보호법을 해석, 적용, 집행하여 개인정보의 처리에 따른 활용을 통한 기술의 개발, 이용과 정보주체의 개인정보를 안전하게 보호에 기여할 수 있기를 기대한다.

.....

주요어 : 개인정보보호법, 가명처리, 암호화조치

학 번 : 2021-28895

목차

Ⅰ. 서론

1. 연구의 목적

가. 4차 산업혁명과 데이터의 이용

인공지능(artificial Intelligence), 기계 학습(machine learning), 사물인터넷(IoT, Internet of Things)과 같은 다양한 기술, 개념이 처음 소개된 때로부터 긴 시간이 흐르지 않았음에도 우리의 일상 요소요소에서 실현되고 있다.

위와 같은 기술들은 소위 '4차 산업혁명'[1]의 개념을 설명할 때 예로 제시되는 것들인데, 실제 우리의 일상에서 구현될 때에는 개인의 위치, 금융거래내역, 맥박과 같은 생체정보 등과 같이 사람과 관련된 정보가 디지털 데이터(digital data)의 형식으로 수집, 저장, 처리되는 형태를 취할 때가 많다.

예를 들어 다수의 사용자 위치를 수집, 처리하여 특정 시간대에 가장 최적화된 이동경로를 찾아주는 '내비게이션(navigation)' 프로그램, 개인의 연령, 직업, 거주지와 그들의 금융거래내역을 비교, 분석하는 등의 처리를 통한 금융상품의 개발 프로그램, 사용자의 맥박, 체온 등을 자동으로 감지하여 냉난방기기 등을 조절하는 프로그램 등이 있고, 이는 이미 상용화된 것들이다.

이와 관련된 새로운 기술을 개발하고 이를 이용하는 과정에서 위와 같은 '사람과 관련된 정보'를 포함한 데이터가 빈번하게 사용되는 것은 필연적이다. 기계 학습을 이용한 암 진단 알고리즘을 생각해 본다면 ① 다양한 병변의 사

---

1) 제네바대 경제학 명예교수이자 세계경제포럼(World Economic Forum, WEF, 일명 '다보스 포럼')의 설립자이기도 한 클라우스 슈밥(Klaus Schwab)이 미국의 국제 관계에 관한 간행물인 포린 어페어스(Foreign Affairs)에 2015. 12. 15.자로 기고한 글에서 소개한 개념으로 2016. 4. 위 포럼에서 상세히 발표되었다. [foreignaffairs.com/articles/2015-12-12/fourth-industrial-revolution; forbes.com/sites/bernardmarr/2016/04/05/why-everyone-must-get-ready-for-4th-industrial-revolution (각 2021. 11. 8. 확인)]

진, 영상이 대량으로 제공되어야만 진단의 정확도를 제고할 수 있을 것이고, ② 이 때 위와 같이 병변에 직접 관련된 정보 외에도 각 병변을 가지고 있던 사람, 즉, 정보주체의 기본적인 정보(성별, 연령, 거주지, 혈액형, 가족의 인적 사항 가족의 암 발병이력 등)도 함께 제공되어야 알고리즘의 효과적인 구현, 검증, 사용이 가능할 것이기 때문이다.

결국 데이터, 그것도 대량의 데이터, 즉 '빅데이터(big data)'를 효과적이고도 창의적으로 처리하는 것은 4차 산업혁명의 시대를 대표하는 각종 기술의 개발에 필수적 전제라고 하겠다.

나. 헌법적 가치의 조화를 통한 '적정한' 데이터 이용의 보장

1) 개인정보자기결정권과 개인정보보호법의 기본 원칙

위에서 예시로 든 여러 기술의 개발, 이용 과정에서 제공되는 데이터의 상당 부분은 사람과 관련된 정보로, 이와 같은 정보 중 일부는 그 내용이나 정보주체와의 관련성, 처리 형태 등의 요소에 따라서는 현재 우리 법체계에서 개인정보의 보호에 관한 일반법인 개인정보보호법이 보호하고 있는 개인정보에 포섭될 수 있을 것이다.

예를 들어 차량의 등록번호, 이용자의 성명, 주민등록번호, 주소, 가족관계, 성별, 연락, 차량의 위치 등은 그 자체로 개인정보이거나, 서로 또는 그 외의 관련 정보와 결합되었을 때 개인정보가 될 수 있다.

헌법재판소는 헌법 제10조(인간의 존엄과 가치, 행복추구권), 제17조(사생활의 비밀과 자유) 및 각종 헌법원리로부터 도출해 낸 개인정보자기결정권을 헌법상 기본권으로 확인하면서, 이를 「자신에 관한 정보가 언제 누구에게 어느 범위까지 알려지고 또 이용되도록 할 것인지를 그 정보주체가 스스로 결정할 수 있는 권리, 즉 정보주체가 개인정보의 공개와 이용에 관하여 스스로 결정할 권리」라고 설명하였다.[2)]

---

2) 헌법재판소 2005. 5. 26.자 99헌마513호 결정

이에 대하여 우리나라의 개인정보보호 관련 기본법인 개인정보보호법은 2011. 9. 30. 시행된 이래 현재까지도, 개인정보의 처리 및 보호에 관한 사항을 정하는 것을 목적으로 천명(법 제1조)하면서, 정보주체가 개인정보의 처리에 관한 정보를 제공 받을 권리, 개인정보의 처리에 관한 동의 여부, 범위 등을 선택하고 결정할 권리 등과 관련된 규정(법 제4조 제1호, 제2호)을 사실상의 기본조항으로서 유지하고 있어, 일응 헌법상 권리인 개인정보자기결정권을 보장하는 입법이라고 볼 수 있다.

위와 같은 헌법재판소 결정례, 개인정보보호법의 목적 및 기본 규정을 종합하면, 최소한 개인정보 보호의 측면에서는 개인정보보호법이 정보주체의 개인정보보호에 방점을 두고, 개인정보처리자가 개인정보를 처리하고자 할 때에는 정보주체의 동의를 원칙적 요건으로 삼고 있다고 할 수 있다.[3)]

2) 4차 산업혁명 시대의 재산권 보장, 학문, 예술, 언론·출판의 자유

그러나 개인정보를 처리하여 새로운 정보를 생성하고, 이를 이용하여 산업활동을 하거나, 이와 같은 과정을 학문적 연구의 대상으로 삼는 등 각종 의사표현에 사용하는 정보주체 이외의 국민(이를 본 연구에서는 '정보이용 주체'라고 하겠다.)에게도 헌법이 보장하는 재산권 행사의 자유, 학문, 예술, 언론·출판의 자유가 있음은 자명하다.

한편, 4차 산업혁명 시대에 대두되는 각종 기술은 대량의 디지털데이터의 빈번한 처리를 전제로 하고 있는 것[4)]으로 기술의 개발, 검증, 구현, 이용의 전

---

3) 개인정보자기결정권은 기본적으로 대국가적 공권이어서 정보주체가 자신의 정보 이용을 하고자 하는 사인에게 자신의 동의를 요구할 권리가 당연히 개인정보자기결정권에서 유래되거나 동일하게 볼 수는 없고, 그러한 동의가 개인정보자기결정권의 본질, 핵심내용에는 해당된다고 보기 어려우므로, '개인정보보호권'이라는 용어가 권리의 내용을 잘 반영한다고 설명하는 견해는 경청할 필요가 있다.(김송옥, 가명정보의 안전한 처리와 합리적 이용을 위한 균형점 -데이터 3법에 대한 헌법적 평가를 겸하여 -, 공법연구 제49집 제2호, p.388) 이에 대하여 정보주체의 동의 없이 가명정보로 처리할 수 있게 된 신법이 과잉금지원칙상 침해최소성, 법익균형성을 위반하여 개인정보자기결정권을 침해한 입법이라는 주장이 있는데(김희정, 가명정보 미동의 처리의 기본권 침해 검토, 법학논총 제45권 제1호, p.58-59), 입법으로 인한 기본권 침해의 기본 요건을 충족하였다고 보기 어려우므로 동의하기는 어렵다.

4) 이미 법은 개인정보의 처리를 「개인정보의 수집, 생성, 연계, 연동, 기록, 저장, 보유, 가공,

단계에서 기본 법체계의 원칙인 정보주체의 동의를 완벽히 관철하는 것은 현실적으로 불가능하거나, 상당한 장애가 될 것이다.

그런데 개인정보보호법은 정보이용 주체들에게 개인정보를 처리, 이용할 때, 당해 정보의 정보주체로부터의 동의를 원칙적 요건으로 삼고 있으므로 이러한 면에서는 위 법이 정보이용 주체들이 향유하는 각종 헌법상 권리를 제한하는 측면이 있다고 하겠다.

이와 같은 상황에서 입법자는 2020. 2. 4. 소위 데이터 3법인 개인정보보호법, 정보통신망이용촉진및정보보호등에관한법률, 신용정보의이용및보호에관한법률을 개정하여 각 신법이 2020. 8. 5.부터 시행(이하에서는 본문에 개인정보보호법의 개정 법률을 지칭할 때에는 '신법'으로, 각 조문을 인용할 때는 '법'으로 각 약칭한다.)되고 있다.

뒤에서 상세히 살펴보겠으나, 데이터 3법의 주된 개정 이유는 '4차 산업혁명 시대를 맞아 핵심 자원인 데이터의 이용 활성화를 통한 사회적 규범을 정립하는 것'에 있다고 요약할 수 있는데[5], 이는 개인정보의 적정한 이용을 활성화하여 기술 발전을 유도하고 기업활동에 활력을 주어야 한다는 사회적 요구가 입법에 반영된 것[6]으로 평가할 수 있다.

3) 개인정보자기결정권과 재산권 등의 조화로운 실현

4차 산업혁명의 개념을 역설하는 입장에서도 위와 같은 디지털 데이터의 이용으로 신산업이 발전하고 새로운 형태의 재화, 자산이 축적될 것임을 전망하는 한편, 디지털 데이터의 오용에 따른 사생활, 사유재산, 거버넌스 등에 대한 침해 우려가 증가되고 있어 각자의 개인정보가 적정히 관리, 보호되고 있다는 믿음을 사회구성원들에게 심어주는 것이 중요한 과제라고 보고 있는데[7], 이

편집, 검색, 출력, 정정(訂正), 복구, 이용, 제공, 공개, 파기(破棄), 그 밖에 이와 유사한 행위」로 정의하여 개인정보에 대한 직접적인 사용 행위의 전후 연결된 행위, 필수적으로 관련된 행위 일체를 '처리'로서 법의 규율 대상으로 삼고 있다. (법 제2조 제2호)

5) https://www.law.go.kr/법령/개인정보보호법, 법률 제16930호 제정·개정 이유(2022. 10. 24. 확인)

6) 개인정보 보호규율이 지나치게 엄격해서 4차 산업혁명시대의 핵심 산업인 빅데이터, 인공지능 등 데이터 이용을 근간으로 한 신산업의 발전이 지체되고 있다는 비판도 제기된다.(김현숙, 과학적 연구목적을 위한 개인정보 처리에 관한 비교법적 연구, 정보법학 제24권 제1호, p.117)

역시 본 연구도 함께 하고 있는 문제의식이다.

위와 같이 개정된 신법에서 「개인정보의 적정한, 적법한 이용」이 명시적으로 선언되지는 않았지만, 개정 내용의 상당 부분이 개인정보의 처리에 관한 것인 점, 데이터를 활용한 자유로운 기술의 개발, 이용을 위해서는 개인정보에 포함되는 데이터의 처리가 필수적이라는 고려에 따른 개정인 점 등을 고려할 때, 정보이용 주체의 개인정보 처리에 관한 재산권 기타 기본권의 보장과 정보주체의 개인정보자기결정권의 보호가 어떻게 조화될 수 있을 것인지의 문제[8]는 현 시점 우리 정보보호법제의 해석, 적용 등 운용에 있어 가장 주요한 과제[9]라고 할 수 있다.

본 연구에서는 적정하고도 안전한 개인정보의 처리를 통하여 사회 구성원 대부분이 각자의 개인정보 처리에 거부감을 가지지 아니하는 것에서 개인정보의 효과적인 처리가 시작된다는 인식, 즉, 사회 구성원들이 개인정보를 안전하게 보호하면서도 이를 적정히 이용하는 것이 가능하다는 공감대를 형성할 필요가 있다는 인식을 가지고 아래와 같이 논의를 진행하고자 한다.

이와 같은 문제의식에서 위와 같은 신법의 개정 취지를 고려할 때 신법에서 개정된 내용 중 주목되는 것은 가명처리, 가명정보이다. 단순히 개인정보를 안전히 관리하고 개인정보의 침해를 억제하고자 한다면 개인정보를 취득한 직후, 목적을 달성하면 즉시 폐기하면 그만일 것이다. 그러나 상당수의 경우는 개인정보를 계속하여 보관하면서 가공, 전송, 생성, 연산 등 다양한 형태로 처리하는 유형에 속할 것이고, 이러한 방식에 따른다면 처리되는 데이터의 양과 내용이 충실할수록 처리된 정보의 가치는 높아질 것이나, 그만큼 정보주체가 식별될 위험성도 높아진다.[10]

---

7) 예를 들어, Alan Marcus, "Data and the fourth industrial revolution", World Economic Forum (weforum.org/agenda/2015/12/data-and-the-fourth-indutrial-revolution, 2021. 11. 8.확인)

8) 이와 관련하여 개인정보자기결정권이 헌법상 확인되었다는 이유만으로 정보주체에 대하여 절대적인 통제권이 부여된 것처럼 오해되고, 그러한 오해가 기본권의 대사인효 논의까지 이어져 민사상 개인정보이용을 제한하는 근거로 이해되어서는 안 된다는 견해(이인호, 「개인정보보호법」 상의 '개인정보' 개념에 대한 해석론, 정보법학 제19권 제1호, p.68)가 있는데 적절한 지적이라고 생각된다.

9) 개인정보보호가 기본권으로 인정되고 있지만 절대적으로 보호되는 것은 아니고 다른 기본권과 비례성의 원칙에 따라 조화를 이루어야 한다는 견해도 있다.[박노형, 개인정보보호법(2020), p.36)]

결국 이와 같은 상황에서 개인정보 침해를 방지하면서 효율적으로 개인정보를 처리하기 위해서 도입된 것이 신법의 가명처리이고, 그에 따라 생성되는 정보가 가명정보라고 생각된다.

한편, 가명처리의 개념에 비추어볼 때 암호화조치도 가명처리의 한 방식이 될 것이고, 구법에서도 제한적이나마 암호화조치가 규정되었으며 아래에서 언급하는 것과 같이 가명처리 중 개인정보의 처리를 통한 이용의 측면에서는 가장 많이 사용될 수밖에 없는 것이 암호화조치이므로 암호화조치의 중요성도 상당하다.

따라서 본 연구에서는 개정된 개인정보보호법이 가명처리, 암호화조치 등에 접근하고 있는 방식을 살펴보고, 외국의 입법례 및 가명처리, 암호화조치의 개념, 적정한 가명처리 등의 요건을 검토하면서 기존 판례의 해석론이나 현행 법령, 가이드라인에서 입법으로 개선할 부분, 해석으로 극복할 부분을 법해석학과 가명처리 기술의 측면에서 검토하는 것을 목적으로 삼고자 한다.

2. 연구의 범위, 방법

본 연구에서는 위와 같이 신법상 도입된 가명처리, 가명처리의 한 방안으로서의 암호화조치에 초점을 맞추어 논의를 진행할 것이다.

우선 신법의 주요 개정 내용 중 본 연구가 관심을 가지고 있는 가명처리, 암호화 조치 관련 부분을 위주로 살펴보면서 개정취지, 구법상의 암호화조치 의무, 시행령, 고시 등 하위법령 및 개인정보보호위원회 등 관계 기관의 가이드라인의 내용을 정리하여 법령의 해석론, 입법상 개선 필요사항을 정리해 보겠다.

그 다음으로는 우리 개인정보보호법제에 큰 영향을 미친 유럽개인정보보호규정(General Data Protection Regulation, 이하 'GDPR'이라 한다.)과 암호화

10) 이와 같이 표현하면서 데이터의 양과 개인정보보호는 트레이드오프(trade-off) 관계에 있다고 설명하는 견해가 적절히 표현한 것으로 보인다.(정영진, 보건의료데이터와 개인정보보호와의 관계에 대한 소고, 법학논총 제34권 제3호, p.217)

기술, 의료 관련 산업이 발달한 미국의 연방법인 의료정보보호법(Health Insurance Portability and Accountability Act of 1996, 이하 'HIPAA'라 한다.)의 제 규정 중, 가명처리의 상위 개념인 비식별조치, 가명처리, 익명정보, 암호화조치 관련 규정을 살펴보아 논의의 바탕으로 삼을 것이다.

그 다음으로는 편을 바꾸어 우선 비식별조치, 가명처리 및 암호화조치의 개념, 요건을 살펴보면서 특히 비식별조치에 관한 해석론, 구법에서부터 규정되어 있었던 암호화조치의 개념, 가명처리로서의 암호화조치의 의미, 각종 암호화 알고리즘에 관한 개념, 특성 등을 기존의 논의를 토대로 개관하고, 그 과정에서 구법에서 제기되었던 논란이나 GDPR의 관련 규정을 다시 상기하여 볼 것이다.

이어서 수리암호를 바탕으로 한 암호화 알고리즘의 특성, 장단점을 살펴볼 것인데, 그 이유는 신법에 따라 개인정보의 안전한 처리를 넘어서 적정한 이용까지 추구하는 경우, 수학의 개념을 이용한 수리암호화가 가명처리 기법으로 빈번히 활용되고 있으나, 기존 판례나 법령, 가이드라인에서는 적정한 암호화조치에 대한 내용을 오해하거나 상세한 규율을 회피하였으므로 각종 암호화 알고리즘의 개념, 장단점 및 응용 사례를 정리하여 암호화조치의 특성에 알맞은 입법상 개선점, 법령의 해석, 적용, 집행에 대한 적절한 기준을 제시할 필요가 있다고 판단하였기 때문이다.

그 다음으로 구법이 시행되던 시기의 각급 법원의 판결례 중에서 본 연구의 범위 내에 있는 주제와 관련된 판결례를 위주로 그 요지, 당사자의 주장, 법원의 판단을 정리한 후, 신법 시행 이후에도 유효한 판결인지, 개인정보의 안전한 보호 및 개인정보의 적정한 이용의 측면에서 비추어볼 때 부당한 판결은 없었는지, 신법을 가정적으로 적용하는 경우에는 어떠한지, 기존 판결례를 통해서 우리 법령 등에 대한 개선점을 찾아낼 수 있는지 등에 대하여 고민해 보고자 한다.

끝으로 이상과 같은 본 연구의 내용을 토대로 입법론, 입법취지에 따른 법해석론 및 기술적인 관점에서의 암호화조치 활용 필요성 등 3가지 측면에서 개선방안을 도출하여 제언하여 볼 것이다.

Ⅱ. 개인정보보호법과 가명처리

1. 2020. 2. 4. 데이터 3법의 개정

2020. 2. 4. 법률 제16930호로 개정된 개인정보보호법의 개정이유는 '4차 산업혁명 시대를 맞아 핵심 자원인 데이터의 이용 활성화를 통한 신산업 육성이 범국가적 과제로 대두'되고 있는 가운데 '신산업 육성을 위해서는 인공지능, 클라우드 등 신기술을 이용한 데이터 이용이 필요'하므로, '안전한 데이터 이용을 위한 사회적 규범 정립이 시급한 상황'이어서, '정보주체의 동의 없이 과학적 연구, 통계작성, 공익적 기록보존 등의 목적으로 가명정보를 이용할 수 있는 근거를 마련하되, 개인정보처리자의 책임성 강화 등 개인정보를 안전하게 보호하기 위한 제도적 장치를 마련하는 한편' 등으로 설명되고 있다.[11)]

그 외 나머지 데이터3법의 개정이유를 살펴보아도, 신용정보법의 경우 '데이터는 사물인터넷, 인공지능 등으로 대표되는 4차 산업혁명의 흐름 속에서 혁신성장의 토대가 될 것으로 기대되고 있는데, 특히, 금융분야에서는 소비·투자 행태, 위험성향 등 개인의 특성을 반영한 맞춤형 금융상품의 개발 등 데이터의 이용가치가 매우 높으나, 우리나라는 빅데이터 이용률이 저조하고, 빅데이터 이용과 분석수준도 다른 나라에 비해 뒤처져 있는 실정인바, 데이터 경제로의 전환이라는 전 세계적 환경변화를 적극적으로 수용하면서 적극적인 데이터 이용으로 소비자 중심의 금융혁신 등의 계기를 마련하기 위하여 빅데이터 분석·이용의 법적 근거를 명확히 함과 동시에, 빅데이터 이용에 따라 발생할 수 있는 부작용을 방지하기 위한 안전장치를 강화'하겠다고 설명하고 있고[12)], 정보통신망법의 경우, '데이터를 핵심 자원으로 하는 4차 산업혁명 시대를 맞아 개인정보의 보호와 이용을 조화시킬 수 있는 제도를 마련하여 더 나은 국민의 삶을 만들어 나가야 할 시점'이라고 표현하였다.[13)]

---

11) 앞의 개정이유

12) https://www.law.go.kr/법령/신용정보의이용및보호에관한법률/(16957,20200204), 법률 제16957호 제정·개정 이유(2022. 10. 24. 확인)

이상을 종합하면, 개인정보보호법 등을 포함한 데이터 3법의 개정은 4차 산업혁명 관련 주요 기술개발, 신산업육성을 위하여 데이터의 적법한 이용을 촉진[14]하기 위하여 데이터 이용에 관한 법적 근거를 마련하면서 그와 같은 데이터 이용에 따른 부작용을 방지하는 것에 그 목적이 있다고 정리할 수 있겠다.

특히 개인정보보호법의 개정 이유는 위와 같은 데이터 이용의 활성화의 주요한 수단으로 정보주체의 동의 없는 개인정보의 이용을 명시적으로 내세우고 있고, 이것이 신법 개정이 가져올 가장 큰 변화라고 생각된다.[15]

2. 구법상 암호화조치의무

구법에서는 신법이 도입한 가명처리, 가명정보의 개념이 규정되지 아니하였으나, 개인정보의 정의 규정(구법 제2조 제1호의 「개인을 알아볼 수 있는(해당 정보만으로는 특정 개인을 알아볼 수 없더라도)」)과 개인정보의 목적 외 이용・제공 규정(구법 제18조 제2항 제4호의 「4. 통계작성 및 학술연구 등의 목적을 위하여 필요한 경우로서 특정 개인을 알아볼 수 없는 형태로 개인정보를 제공하는 경우」)의 해석상 정보주체의 '식별가능성'이 요건으로 도출될 수밖에 없었다.(이상 강조점은 연구자가 부여하였다.)

이에 따라 처리 대상 정보의 식별가능성을 없애거나 낮추는 조치인 비식별조치(de-identification)가 당시 주무부처인 행정자치부의 가이드라인에서 요구

---

13) https://www.law.go.kr/법령/정보통신망이용촉진및정보보호등에관한법률/(16955,20200204) 법률 제16955호 제정·개정 이유(2022. 10. 24. 확인), 다만 정보통신망법의 경우 아래에서 살펴보는 것과 같이 위 개정을 통하여 개인정보보호법과 유사, 중복되는 조항을 정비하고자 개인정보보호에 관한 사항이 삭제되었다.

14) 개보위에서 제작, 공표한 가명정보 처리 가이드라인에서는 신법의 개정배경을 설명하면서 "4차 산업혁명 시대 신성장 동력인 데이터 활용에 대한 시대적 요구를 반영한 것으로, 개인정보처리자가 통계작성, 과학적 연구, 공익적 기록보존 등을 위한 목적으로 개인정보를 가명처리하여 활용할 수 있는 기반이 새롭게 마련되었다."고 정리한다.[개보위, 가명정보 처리 가이드라인(2022), p.5]

15) 이에 대하여는 정보주체의 동의가 전제된 개인정보 활용의 예외를 인정하여 개인정보의 활용과 보호라는 두 가지 목적을 함께 달성할 수 있는 대안으로, 이른바 '비식별조치'에 대한 논의가 확산, 발전된 결과라고 해석한 견해[고학수 등 7인, 인공지능 시대의 개인정보 보호법(2022. 5.) p.5)]가 있는데, 적절한 분석이라 생각된다.

되었는데[16], 이는 뒤에서 살펴보는 것과 같이 가이드라인이 법현실을 규율하는 것에는 명백한 한계가 있었고, 이는 신법 개정에 하나의 계기를 제공하였다고 볼 수 있다.

그러나 구법에서도 암호화조치가 법적 의무로 부과된 부분이 있었고, 현재 신법에서도 그대로 유지되고 있어 간략히 살펴보겠다.

우선, 구법 제24조는 「고유식별정보의 처리 제한」이라는 표제 하에 제3항에서, "개인정보처리자가 제1항 각 호에 따라 고유식별정보를 처리하는 경우에는 그 고유식별정보가 분실·도난·유출·위조·변조 또는 훼손되지 아니하도록 대통령령으로 정하는 바에 따라 암호화 등 안전성 확보에 필요한 조치를 하여야 한다."고 규정하여, 주민등록번호, 여권번호, 면허번호, 외국인등록번호(구법 시행령 제19조 각호)와 같은 고유식별정보에 대하여는 안전성 확보에 필요한 조치를 하도록 개인정보처리자에 대하여 의무를 부과하면서 '암호화'를 그 조치의 예시로 들고 있었고. 특히 고유식별정보 중 주민등록번호에 대하여는 구법 제24조의2 제2항이 암호화조치를 반드시 취할 것을 요구하였다.

이상의 규정은 신법의 개정 당시에도 그대로 유지되었고, 각 법률규정에 대한 시행령도 변동이 없으며. 구법 체제 하에서는 행정안전부장관의 고시로서 법령에서 위임된 내용을 정하던 개인정보의 안전성 확보조치 기준(당시 행정안전부고시 제2019-47호, 현재 폐지)도 현재 주무 관청만 바뀐 채 동명의 고시로 그대로 유지되고 있는데[17], 신법이 대폭 개정된 상황에서도 고시에 특별한 변경이 없으므로 야기되는 문제점은 아래에서 살펴볼 필요가 있다.

3. 신법의 주요 개정 내용

개정 개인정보보호법의 주요내용 중 본 연구에서 관심을 가지고 있는 내용 위주로 살펴보면 아래와 같다.

---

16) 고학수 등 7명, 앞의 책, p.23

17) 바이오정보라는 표현 대신 생체정보라는 표현으로 변경되었고, 생체정보 중 인증 기능을 추가로 규율한 외에는 개정된 내용이 없다.

가. 개인정보의 정의 확대, 가명처리의 근거 마련

우선 구법에서 개인정보의 정의(구법 제2조 제1호)에 포함되었던 '성명, 주민등록번호 및 영상 등을 통하여 개인을 알아볼 수 있는 정보'를[18] 신법에서는 첫 번째 유형으로 분리하고(제2조 제1호 가목), 구법에서 '해당 정보만으로는 특정 개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 알아볼 수 있는 것을 포함한다.'고 부기되었던 부분은 신법에서 제2조 제1호 나목으로 분리되어 「해당 정보만으로는 특정 개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 알아볼 수 있는 정보. 이 경우 쉽게 결합할 수 있는지 여부는 다른 정보의 입수 가능성 등 개인을 알아보는 데 소요되는 시간, 비용, 기술 등을 합리적으로 고려하여야 한다.」고 정하여 별개의 개인정보 유형으로 특정하면서 그 판단 기준을 제시하는 진일보한 입법 기술을 보여주고 있다.

특히 신법에서는 개인정보의 일부를 삭제하거나 일부 또는 전부를 대체하는 등의 방법으로 추가 정보가 없이는 특정개인을 알아볼 수 없도록 처리하는 것을 가명처리로 정의하여 가명처리의 법적 근거를 마련하였고(제2조 제1호의2를 신설), 아래에서 살펴보는 것과 같이 가명처리의 경우 일정한 특례를 정하게 되었다.[19]

---

18) 강학상 이와 같은 정보를 그 자체로 정보주체가 식별될 수 있는 것이라고 보아 식별자(identifier)라고 칭하기도 하는데, GDPR이나 아래에서 살펴보는 미국 연방 법률인 의료정보보호법(HIPAA)에서도 사용되는 표현이다. 한편 위 가목의 정보를 '개인식별정보'라고 표현하는 방식도 확인된다.(고학수 등 7인, 앞의 책, p.10) 아래 같은 호 나목에서 정한 정보를 식별자에 대비되는 준식별자(quasi-identifier)라고 일컫는 수밖에 없다면 준식별자 대신 '개인식별가능정보'로 새기는 것(고학수 등 7인, 앞의 책, p.12)이 가능할 것이므로, 가목에 해당하는 정보를 '개인식별정보', 나목에 해당하는 정보를 '개인식별가능정보'로 칭하는 것이 직관적 이해에 도움을 줄 것으로 보인다. 실제 법 제24조에는 주민등록번호, 여권번호 등 정보주체 마다 부여되어 일대일 대응이 가능한 정보를 '고유식별정보'로 분류하고 있으므로 위와 같은 용법이 상당하다고 판단한다.

19) 한편 고학수 등 7인, 앞의 책, p.16에서는 법 제23조의 민감정보(사상, 신념, 노동조합, 정당의 가입 등), 법 제24조의 고유식별정보의 경우 가명처리를 할 수 있는지 여부에 대하여 법이 명확하게 규율하지 않고 있고, 개보위의 가명처리 가이드라인에서 언급하고 있다고 소개하고 있다. 위 가이드라인은 민감정보, 고유식별정보도 가명처리가 가능하나 고유식별정보 중 주민등록번호는 불가능하다고 설명한다.(개보위, 위 가이드라인, p.14) 아래에서 상세히 논증하는 것과 같이 '가이드라인'은 법의 취지를 예를 들어 설명하는 것이어서 사례형 유권해석에 해당할 뿐, 규범력이 있다고 보기 어려우므로 법령에 정해져야 할 규범적

그러나 가명처리의 개념을 도입하면서도 안전한 가명처리의 방식이나 가명처리의 결과물이 '특정개인을 알아볼 수 없도록 처리된 것'인지 여부를 판단하는 기준(즉, 적정한 가명처리의 판단 기준), 추가 정보의 의미, 요건[20] 등에 관하여 시행령에 위임하는 등의 방법으로라도 규율하지 아니하였고, 이는 아래에서 언급하는 가명처리의 특례 규정에서도 마찬가지이다.

나아가 신법은 위 가목 및 나목에서 정한 정보를 위와 같이 가명처리함으로써 원래의 상태로 복원하기 위한 추가 정보의 사용, 결합 없이는 특정 개인을 알아볼 수 없는 정보, 즉, 가명정보를 개인정보의 범위에 포함하고 있고(제2조 제1호 다목을 신설), 개인정보처리자에 대하여 익명, 가명으로 처리하여도 개인정보 수집목적을 달성할 수 있는 경우, 익명처리가 가능하다면 익명처리, 익명처리로 목적을 달성할 수 없는 경우 가명처리를 할 것을 요구하고 있다.(제3조 제7항)[21]

나. 정보주체의 동의 없는 개인정보의 이용, 제공(법 제15조 제3항 등)

법 제15조 제3항, 제17조 제2항은 개인정보처리자가 「당초 수집 목적과 합리적으로 관련된 범위에서 정보주체에게 불이익이 발생하는지 여부, 암호화 등 안전성 확보에 필요한 조치를 하였는지 여부 등을 고려하여 대통령령이 정하는 바에 따라 정보주체의 동의 없이 개인정보를 이용, 제공」(이하에서는 '법 제15조 제3항 등에 따른 이용'으로 약칭한다.)할 수 있도록 규정하였다.

다만, 아래에서 상론하다시피 수집 목적과의 합리적 관련성이라는 요건이 개인정보처리자의 관점에서는 상당한 제약으로 기능하고 있고, 결국은 시행령

사항을 가이드라인으로 제시하는 것은 부적절하다고 본다. 다만, 위 문제에 대하여는 민감정보, 고유식별정보도 개인정보에 포섭되고, 법 제23조 제1항 제2호, 법 제24조 제1항 제2호에서 '법령에서 처리를 허용하는 경우'를 두고 있으므로 가명처리 또한 이에 해당한다고 해석할 수 있을 것으로 보인다. 이는 아래에서 각종 가이드라인의 문제점을 지적할 때 상세히 논하도록 하겠다.

20) 특히 '추가 정보'에 관하여 구체적인 규율이 없어 명확성의 원칙에 반한다는 견해도 있다. (김송옥, 앞의 논문, p.390)

21) 구법에서는 "개인정보처리자는 개인정보의 익명처리가 가능한 경우에는 익명에 의하여 처리될 수 있도록 하여야 한다."고 규정한 바 있다.(구법 제3조 제7항)

에 의하여 가명처리, 암호화조치를 취하거나, 이에 준하는 정도의 안전성 확보를 위한 조치가 필요하므로 개인정보이용의 활성화라는 목표에 비추어 보면 개정의 의미는 제한적이다.

다. 가명정보의 처리 등에 대한 특례(법 제28조의2 이하 신설)

신법은 앞서 살펴본 가명정보에 대하여 개인정보처리자가 통계작성, 과학적 연구, 공익적 기록보존 등을 위하여 정보주체의 동의 없이 가명정보를 처리할 수 있도록 하되, 서로 다른 개인정보처리자 간의 가명정보의 결합은 개보위 또는 관계 중앙행정기관의 장이 지정하는 전문기관이 수행하도록 하면서(법 제28조의2, 제28조의3 신설), 가명정보를 처리하는 경우 해당 정보가 분실, 도난, 유출, 위조, 변조 또는 훼손되지 않도록 대통령령이 정하는 바에 따라 안전성 확보에 필요한 기술적, 관리적 및 물리적 조치를 하도록 하였다.(법 제28조의4 신설) 법 제28조의4에서는 일단 구체적인 사항을 대통령령에 위임함으로써 법률 규정 자체로는 위에서 언급한 가명정보, 가명처리에 대하여 지적할 수 있는 문제점이 다소 완화 되었으나, 후술하는 것과 같이 대통령령 이하에서도 구체적인 규율이 이뤄지지 않고 있다.

한편, 구법에서는 개인정보의 목적 외 이용, 제공이 예외적으로 가능한 경우로서 통계작성 및 학술연구 등의 목적을 위하여 필요한 경우로서 특정 개인을 알아볼 수 없는 형태로 개인정보를 제공하는 경우를 정한 바 있는데(구법 제18조 제2항 제4호), 위와 같은 '특정 개인을 알아볼 수 없는 형태의 개인정보'가 문언대로라면 익명정보로 볼 여지가 크다. 결국 구법의 위 조항에 따르면 특정한 목적을 위하여 익명정보에 이를 정도로 비식별조치를 한 후에야 개인정보의 수집, 처리 목적 외의 이용, 제공이 가능하다고 볼 수밖에 없어 개인정보의 이용의 측면에서는 한계가 분명하였다.

그러나 신법에서는 '과학적 연구'를[22] 위와 같이 가명처리의 목적으로 명기

22) 신법 제2조 제8호는 과학적 연구를 「기술의 개발과 실증, 기초연구, 응용연구 및 민간투자 연구 등 과학적 방법을 적용하는 연구」로 정의하고 있어, 신법에서는 민간에서 영리를 목적으로 한 연구를 가명처리 특례의 적용에서 배제할 이유가 없게 된다. 같은 취지로 위 연구를 학술적 활동, 학술연구로 제한하는 소극적 해석에 국한하기 어렵다는 견해도

하였으므로 익명정보에 비하여 이용가치가 높은 가명정보를 과학적 연구의 목적으로 사용할 수 있도록 개정한 것이어서 개인정보의 이용을 활성화하는 것에 방점을 찍었다고 할 수 있다.

물론 개인정보보호법에서는 가명정보를 이용하지 않더라도 개인정보를 이용할 수 있는 대안이 있기는 하다. ① 정보주체의 동의를 받는 경우에 개인정보의 이용이 가능함은 자명하고, ② 아래에서 살펴보는 익명정보(법 제58조의2 신설)의 경우 개인정보보호법의 적용에서 배제되므로 특별한 제한이 없을 것이며, ③ 앞서 언급한 법 제15조 제3항 등에 따른 이용의 경우에도 정보주체의 동의 없이도 개인정보를 이용할 수는 있기 때문이다.

그러나 개인정보를 디지털 데이터의 형태로 처리하는 각종 기술은 대개 대량의 데이터를 빈번하게 처리하는 방식을 전제하고 있으므로, 이를 염두에 두고 살펴본다면 위와 같은 대안의 한계는 아래와 같이 명확하다.

① 우선 각각의 데이터의 처리 마다 정보주체의 동의를 받는 것은 그 자체로 지난한 일인데다가, 가사 동의를 받았다고 하더라도 그 시적, 물적, 인적 범위를 명확히 하는 경우 동의의 범위가 지나치게 협소해지고, 반대로 그러한 범위가 명확하지 않다면 동의의 범위에 관하여 해석이 구구해질 것이다.

② 한편, 누구인지 알아 볼 수 없는 완전한 익명정보는 익명처리 과정에서 이용 가치가 낮아질 가능성이 높다.

③ 그리고 법 제15조 제3항 등에 따른 이용의 경우, 수집 목적과의 관련성을 제한[23]으로 두는 이상, 관련성의 제한을 엄격히 볼 경우, 이용가능성이 희박하게 될 것이고, 관련성의 제한을 느슨하게 여길 경우 정보주체의 저항이 불가피할 것이어서, 이와 같은 상황에서 개인정보처리자들이 적극적으로 이를 활용할 것을 기대하기는 어렵고[예를 들어 개인에게 특정한 서비스를 제공하는 회사가 고객의 불만 사례를 대량으로 보관하던 중, 이를 업무 처리 방식 개선에 사용한 경우, 직원의 고객 응대 방식을 점검, 평가하는 소프트웨어에

있다.(고학수 등 7인, 앞의 책, p.78) 위 가명정보처리 가이드라인에서는 새로운 개술, 제품, 서비스 개발 및 실증을 위한 산업적 연구도 가능하다고 보는데(개보위, 위 가이드라인, p.12), 위와 같이 이미 법률에서 '민간투자 연구'를 포섭한 이상 당연한 해석이라 하겠다.

23) 김현숙, 앞의 논문, p.117에서는 개인정보의 수집 시점에서 목적을 특정할 수 없는 경우가 다수 있다고도 지적한다.

고객의 불만 사례를 기계 학습용으로 제공한 경우를 상정해보면, 이러한 개인정보의 이용이 수집 목적(해당 고객과의 분쟁 방지, 해당 고객에 대한 서비스 지원)과 합리적인 관련성이 있는지 여부에 관하여는 구체적 사안에 따라 결론을 달리 할 것이고, 가사 합리적인 관련성을 인정할 여지가 있다고 하더라도 개인정보처리자가 고객과의 법적 분쟁을 우려하게 된다면 이를 쉽게 추진하기는 어려울 것이다.], 무엇보다도 이러한 이용에 대하여도 아래에서 보는 것과 같이 대통령령에서 가명처리 또는 암호화를 요구하고 있어 가명처리에 의한 개인정보의 이용과 구분할 실익이 크지 않다고 본다.

라. 익명정보(법 제58조의2) 등 기타 개정사항

한편 법 제58조의2는 「시간·비용·기술 등을 합리적으로 고려할 때 다른 정보를 사용하여도 더 이상 개인을 알아볼 수 없는 정보」를 개인정보보호법의 적용에서 제외하고 있는데, 표제에서 '익명정보'라는 표현을 명시적으로 사용하지는 않았지만 이를 익명정보로 보는 견해가 다수이다.[24]

위와 같은 개정으로 인하여 구법 당시 개인을 식별할 수 없는 경우(익명정보)와 추가적으로 정보를 사용하면 재식별이 가능한 경우(가명정보)의 구분이 필요하다는 지적[25]을 해소하여 재식별가능성이 없는 경우는 개인정보보호법의 적용을 제외함으로써 합리적인 범위 내에서 개인의 정보, 그러나 익명처리된 정보를 사용할 수 있게 한 입법조치라고 평가할 수는 있다.

그러나 위 조항이 '시간·비용·기술 등을 합리적으로 고려할 때 다른 정보를 사용하여도 더 이상 개인을 알아볼 수 없는 정보'를 사실상 익명정보로 규정하였다고 하나, '다른 정보'가 어떠한 것인지, '개인을 알아볼 수 없는' 수준이 어떠한지는 여전히 법률의 해석으로는 쉽게 답을 할 수 없다고 볼 수 있는데, 이를 가리켜 어느 정도의 비식별조치가 이루어져야 익명정보로 인정될 수 있을지 그 기준이 불분명하다는 지적이 있고[26], 가명처리의 경우에도 개선이

24) 박노형, 앞의 책, p.58, 고학수 등 7명, 앞의 책, p.18
25) 예를 들어, 오길영, 데이터 비식별화 정책에 대한 규범적 비판, 공법연구 제46집 제2호, p.342
26) 고학수 등 7명, 앞의 책, p.19

필요한 지점이라 하겠다.

마. 검토

이상 살펴본 것과 같이 신법이 가명처리, 가명정보의 개념을 도입한 것은 평가받아야 할 것이나, 가명처리를 도입하면서도 적정한 가명처리의 개념, 방식, 적정한 가명처리의 판단 기준, 추가 정보의 개념, 요건, 익명정보의 판단기준에 대하여 기본적인 내용을 제시하지 아니하고 대통령령에 위임하지도 아니하였다는 문제점이 명백하다.[27]

위와 같은 입법조치상의 흠결로 인하여 신법이 개정, 시행되었음에도 법률은 물론 시행령 이하 하위 법규에서도 구체적인 규율을 회피한 탓에 결국 구법 당시와 마찬가지로 법적 근거가 부족하고 규범력이 희박한 가이드라인이 여전히 가명처리 등의 개념을 다루고 있어 수범자들에게 각자의 행위가 적법한 것인지 여부를 예측하기 어려운 실정이다.

물론 이와 같은 입법은 개인정보의 보호, 이용에 관한 기술 발전 속도가 빠르고, 규제대상이 가변적이며 사안별로 판단할 필요가 있다는 고려에서 나온 입법기술의 결과라고 평가할 여지도 있다. 한편으로는 국가가 법령 등으로 구체적인 기술의 예를 들어 이를 권장하는 등의 상세한 규율을 하는 경우, 입법이 현실을 따르지 못하거나, 기술 발전을 저해한다는 등으로 '기술중립성의 원칙'을 해친다는 비판이 제기될 수 있다.

그러나 아래에서 논의하는 것과 같이 규범력을 인정할 수 있는 가장 하위 법규인 고시의 단계에서 조차 적정한 가명처리 등에 대한 판단기준을 제시하지 아니한 가운데 대부분의 행동 준거를 가이드라인에서 찾아야 하는 현행 정보보호법제는 사회구성원들에게 예측가능성을 포함한 법적 안정성을 부여하지 못하고, 경제주체들이 데이터 활용에 보수적인 자세를 취하게 하는 하나의 원인을 제공하고 있다는 것이 본 연구의 문제의식 중 하나이다.

---

27) 예를 들어 가명정보의 처리 목적인 「통계작성, 과학적 연구, 공익적 기록보존」의 해석도 불분명하다는 견해가 있다.(조성훈, 개인정보보호와 형사책임 : 가명정보 특례와 목적의 합리적 관련성을 중심으로, 법학평론 제12권, p.246)

4. 하위 법규의 내용

현행 개인정보보호법 하의 시행령 이하 하위 법규는 모법인 개인정보보호법의 각종 위임을 받은 개인정보보호법 시행령이 대통령령 제32813호로 시행(2022. 10. 20.자) 중이고, 개보위가 위 법률과 시행령의 위임을 받아 신법 시행 이후 제정한 각종 고시들이 세부적인 사항을 규율하고 있다.

이하에서는 개인정보보호법의 제 규정 중, 본 연구에서 관심을 두고 있는 내용 위주로 시행령과 고시의 규정사항을 연결하여 정리하였는데, 법률 단계에서 노정된 문제점이 반복되는 아쉬움이 있다.

가. 법 시행령

1) 정보주체의 동의 없는 개인정보의 추가적인 이용·제공(영 제14조의 2)

영 제14조의2는 법 제15조 제3항 등에 따른 이용의 적법 요건을 판단하는 기준을 제시하고 있는데, ① 수집 목적과의 관련성, ② 개인정보를 수집한 정황, 처리 관행 등에 따라 개인정보의 위와 같은 이용에 관한 예측 가능성의 유무, ③ 정보주체의 이익을 부당하게 침해하는지 여부, ④ 가명처리 또는 암호화 등 안전성 확보에 필요한 조치를 하였는지 여부가 그것이다. 결국 법에서 규정하지 않은 '가명처리'가 하나의 판단기준으로 편입되어 있는데 가명처리야말로 개인정보보호의 유효한 수단이므로 당연한 귀결로 보인다.

다만, 가명처리의 여러 수단 중 하나가 암호화라고 본다면 위와 같이 '가명처리 또는 암호화'라는 규정형식은 마치 가명처리와 암호화는 구별되는 개념인 것처럼 오해될 여지가 있다.

또한, 여러 가지 검토 요건 중 하나로 가명처리 또는 암호화를 들고는 있으나 나머지 요건이 모두 충족되었을 때 가명처리나 암호화조치가 취하지 아니한 상태의 개인정보를 그대로 이용하는 것을 상정하기란 어려운 일이어서 결

국은 가명처리를 하여야 이용 가능한 것이 아니냐는 의문이 제기될 수 있다.

덧붙여 위 조항에서도 적정한 가명처리나 암호화가 어떠한 것이어야 하는지에 대하여는 일응의 기준 조차 제시되지 않고 있고, 다른 조항과 다르게 고시 등 하위 법규의 더 이상의 위임도 하지 않고 있어서 나머지 세부사항은 여타 법령의 가명처리, 암호화조치 관련 내용으로 미루어 짐작할 수밖에 없는 실정이다.

2) 고유식별정보의 처리 제한(영 제21조 → 영 제30조, 제48조의2)

고유식별정보의 처리 제한에 관한 법 제24조 제1항은 고유식별정보에 관하여 대통령령에 위임하고 있는데 이에 따른 영 제19조는 주민등록번호, 여권번호, 운전면허번호, 외국인등록번호를 고유식별정보로 정하고 있다.

이어서 고유식별정보의 처리 시 암호화 등 안전성 확보에 필요한 조치를 요구하고 있는 법 제24조 제3항으로부터 위임받은 영 제21조는 법에서 위임한 안전성 확보에 관한 조치의 상세 내용을 아래에서 정리하는 영 제30조, 제48조의2를 준용하고 있다.

한편, 고유식별번호 중 주민등록번호의 경우 법 제24조의2에서 좀 더 강화된 처리 제한을 정하고 있는데, 동 조항에서 위임한 영 제21조의2는 이미 법률에서 규정한 '암호화조치'를 다시 언급하고 있을 뿐, 적정한 암호화조치에 대한 기준을 규율하지 아니한 채, 재차 고시에 위임하고 있다.(같은 조 제3항)

그 위임을 받은 개보위의 고시가 「개인정보의 안전성 확보조치 기준」인데 위 고시는 영 제21조의2 외에도, 앞서 살펴 본 고유식별정보의 처리 시 안전성 확보조치에 관한 영 제21조 및 아래에서 정리하는 영 제30조의 위임을 받고 있다.

3) 가명정보에 대한 안전조치의무(영 제29조의 5 → 영 제30조, 제48조의2)

법 제28조의4 제1항이 가명정보를 처리하는 경우에 원래의 상태로 복원하기 위한 추가 정보를 별도로 분리하여 보관, 관리하는 등 해당 정보가 분실·도

난·유출 등이 되지 않도록 대통령령으로 정하는 바에 따라 안전성 확보에 필요한 기술적, 관리적 및 물리적 조치를 할 것을 요구하고 있음은 위에서 정리한 것인데, 그러한 조치에 관한 사항을 위임받은 영 제29조의5는 여러 가지의 안전성 확보조치를 열거하면서 영 제30조, 제48조의2에 따른 안전성 확보조치, 가명정보와 추가 정보의 분리보관, 가명정보와 추가 정보에 대한 접근 권한의 분리를 정하고 있다.

4) 개인정보의 안전성 확보 조치(영 제30조)

법 제29조는 개인정보의 이용, 제공과 관련된 규정은 아니나, 개인정보를 보관하고 있는 개인정보처리자에 대하여 개인정보의 분실, 도난, 유출, 위조, 변조, 훼손을 방지하기 위하여 안전성 확보에 필요한 기술적, 관리적 및 물리적 조치를 하여야 함을 규정하면서 그러한 조치의 상세를 영 제30조에 위임하고 있고, 이는 구법의 규정이 그대로 유지되고 있는 것이다.

영 제30조는 고유식별정보, 가명정보에 대한 안전성 조치의무까지 함께 규율하고 있는 조항임은 정리해 보았다. 즉, ① 개인정보의 안전한 처리를 위한 내부 관리계획의 수립·시행, ② 개인정보에 대한 접근 통제 및 접근 권한의 제한 조치, ③ 개인정보를 안전하게 저장·전송할 수 있는 암호화 기술의 적용 또는 이에 상응하는 조치, ④ 개인정보 침해사고 발생에 대응하기 위한 접속기록의 보관 및 위조·변조 방지를 위한 조치, ⑤ 개인정보에 대한 보안프로그램의 설치 및 갱신, ⑥ 개인정보의 안전한 보관을 위한 보관시설의 마련 또는 잠금장치의 설치 등 물리적 조치를 열거하고 있어 개인정보처리자가 취하여야 할 조치를 비교적 상세히 규정하고는 있으나, 개인정보에 대한 접근 통제, 접근 권한의 제한, 암호화 기술, 접속기록의 보관 및 위조, 변조 방지를 위한 조치, 보안프로그램의 내용 등에 관한 세부사항은 마찬가지로 개보위의 고시인 개인정보의 안전성 확보조치 기준에 위임하고 있다.

5) 정보통신서비스 제공자 등에 대한 특례(영 제48조의2)

한편, 영 제48조의2는 정보통신서비스 제공자와 그로부터 이용자의 개인정보를 정보주체의 동의에 따라 제공받은 자에 대하여 영 제30조가 요구하는 조치에서 한층 강화된 안전성 확보 조치의무를 부담하게 하고 있고, 이 조항은 영 제30조와 마찬가지로 고유식별정보, 가명정보에 대한 안전성 확보 조치의무를 구성하기도 한다.

세부적인 조치 사항으로는 개인정보의 안전한 처리를 위한 내부관리계획 수립(제1호), 개인정보에 대한 불법적인 접근을 차단하기 위한 조치(제2호), 접속기록의 위·변조 방지를 위한 조치(제3호), 개인정보가 안전하게 저장·전송될 수 있도록 하기 위한 조치(제4호), 개인정보처리시스템 및 개인정보취급자가 개인정보 처리에 이용하는 정보기기에 컴퓨터바이러스, 스파이웨어 등 악성프로그램의 침투 여부를 점검, 치료할 수 있는 백신소프트웨어 설치 및 주기적 갱신조치(제5호) 등이 규정되어 있고, 세부 기준을 개보위의 고시에 위임하였으니 그것이 「개인정보의 기술적·관리적 보호조치 기준」이다.

특기할 만한 것은 본 조항 제4호가 암호화조치의무에 대하여 비교적 상세히 규율하고 있다는 것인데, 가목에서 비밀번호의 일방향 암호화 저장을 요구하고 있고, 나목에서 주민등록번호, 계좌정보, 일부 민감정보 등에 대하여도 암호화하여 저장할 것을 요구하고 있으며, 라목에서 암호화 기술을 이용한 보안조치를 요구한다는 것이다.

정보통신제공자 등에 대하여 좀 더 강화된 안전성 확보 조치의무를 부과하고 있다는 점, 특정 정보에 대하여는 암호화를 필요적 조치로 정하고 있다는 점 등에서 평가할 수 있는 규정이라 하겠으나, 적정한 '암호화'의 기준을 대강이라도 설정하지 아니하였다는 한계가 지적된다.

6) 소결

이상 법 시행령의 내용을 정리한 바, 아래와 같은 문제점이 정리된다.

우선, 가명처리된 가명정보에 대한 안전성 확보 조치 외에는, 가명처리에 관한 구체적인 규율이 시행령 단계에서도 전무하여 어떠한 가명처리가 적정한 것인지, 적정한 가명처리를 판단하는 기준이 규율되어 있지 아니한데, 이는 이

미 법률의 단계에서 어떠한 위임도 하지 않은 것에 기인한다.

또한, 법 제15조 제3항 등에 따른 이용에 대하여는 법률에서 위임받은 사항, 즉, 각 규정에 따른 정보주체의 동의 없는 이용의 요건을 나름대로 상세히 규율하였으나, 가명처리, 암호화조치를 요건으로 제시하고도 하위 법규인 고시에 위임하지 않은 맹점을 드러내고 있다.

그 외에 시행령은 암호화조치에 관한 위임을 받았음에도 개략적인 규율을 생략한 채 고시에 재위임하고 있는데 이는 위임입법금지 원칙에 합당하지 의문이 제기될 수 있다.[28)]

법 시행령은 법률로부터 구체적이고 직접적인 위임을 받은 사항이 다수 있으므로 법규명령으로서 규범력이 충분히 인정될 수 있을 것임에도, 위임받은 사항에 대하여 불충분하게 규율하거나 이를 재차 고시에 위임하고 있어 법을 해석, 적용, 집행하는 것에 혼란을 야기할 우려가 크다고 생각한다. 법 시행령이 법률과 함께 보완되어야 할 사항은 아래에서 일괄하여 정리하고자 한다.

나. 고시

이미 언급한 것과 같이 법률, 시행령의 일부 규정과 관련된 고시는 개인정보의 안전성 확보조치 기준과 개인정보의 기술적·관리적 보호조치 기준인데 각 2020. 8. 11.부터 적용되고 있다. 아래에서는 본 연구의 논의 범위 내에서 각 고시의 내용을 정리하여 보겠다.

1) 개인정보의 안전성 확보조치 기준

위 고시는 신법이 시행됨에 따라 개보위가 주무관청으로서 법령의 위임을 받아 제정된 것으로 구법 시대와 명칭이 동일하고 내용도 대동소이한데, 신법이 가명정보, 가명처리, 정보주체의 동의 없는 개인정보의 이용, 익명정보의 개념을 도입하였으므로 위 고시도 주무관청이 바뀌면서 새로이 제정되는 이

28) 이러한 점을 가명처리 결합 관련 조항에서 지적하면서 복위임금지의 일반 법리를 위반하였다고 보는 견해도 있다. (김송옥, 앞의 논문, p.393)

상, 신법의 개정취지에 맞게 새로운 규율을 하는 것이 상당하다고 본다.

위 고시에서 본 연구에서 논의하는 내용과 관련 있는 조항은 제7조 '개인정보의 암호화' 부분이다. 위 조항은 개인정보처리자로 하여금, 고유식별정보, 비밀번호, 생체인식정보를 정보통신망을 통하여 송신하거나 보조저장매체 등을 통하여 전달하는 경우(제1항), 비밀번호 및 생체인식정보를 저장하는 경우(제2항), 인터넷 구간 및 인터넷 구간과 내부망의 중간 지점(DMZ : Demilitarized Zone)에 고유식별정보를 저장하는 경우(제3항)에는 암호화조치의무를 필요적으로 규정하고 있는 한편, 내부망에 고유식별정보를 저장하는 경우에는 법 제33조에 따른 개인정보 영향평가의 결과 또는 암호화 미적용 시 위험도 분석 결과에 따라서는 암호화를 하지 않을 수 있도록 하고도 있어(제4항) 의문이 제기될 여지를 남기고 있다.

다만, 비밀번호를 저장하여 보관하게 되는 경우에는 '복호화되지 아니하도록 일방향 암호화'할 것을 요구하고, 내부망에 고유식별정보를 저장하는 경우에는 암호화의 적용여부, 적용범위 등을 정할 수 있도록 하였으며, '안전한 암호화 알고리즘', '상용 암호화 소프트웨어'를 사용할 것을 요구하고 있다.

2) 개인정보의 기술적·관리적 보호조치 기준

위 고시 또한 구법 시대에 이미 방송통신위원회가 정보통신망법의 위임을 받아 제정, 시행하였던 고시[29]와 명칭과 내용이 동일한데, 구법 시대에는 정보통신망법이 정보통신서비스 제공자 등에 대하여 개인정보의 처리 시 분실·도난 등을 방지하고 개인정보 안전성 확보를 위하여 필요한 기술적·관리적 보호조치의 의무를 부과하고 있었기 때문이다.

신법의 개정, 시행에 따라 정보통신망법에서 규율하던 정보통신서비스 제공자에 관한 규정을 개인정보보호법에 추가 신설함에 따라 개보위가 같은 명칭의 고시를 제정하게 되었는데, 개인정보 안전성 확보조치 기준과 마찬가지로 구법 시대의 고시 내용과 큰 차이가 없다.

---

29) 방송통신위원회고시 제2019-13호

고시는 본래 정보통신서비스 제공자에 대하여는 법 제29조의 안전조치의무에 좀 더 무거운 의무를 추가로 부과한 시행령 제48조의2 제3항의 위임을 받은 것이기는 하나, 개인정보 안전성 확보조치 기준의 경우에서 보듯 고유식별정보, 가명정보에 대한 안전성 확보 조치에도 해당된다.

위 고시에도 앞서 살펴본 개인정보 안전성 확보 조치 제7조와 유사한 암호화조치의무 관련 규정이 있으니 제6조가 그것이다.

제6조에서는 비밀번호의 일방향 암호화 저장의무(제1항), 주민등록번호 등 고유식별정보와 신용카드번호, 계좌번호, 생체인식정보의 암호화 저장의무(제2항), 정보통신망을 통한 이용자의 개인정보, 인증정보를 송·수신할 때 안전한 보안서버 구축 조치의무(제3항), 개인정보의 컴퓨터, 모바일 기기 및 보조저장매체의 저장 시 암호화조치의무(제4항)를 부과하고 있어 정보통신서비스 제공자에 대하여는 한층 더 강화된 암호화조치의무를 규정하고 있다. 즉, 일정한 경우에 암호화조치를 취하지 않을 수 있는 여지를 남기고 있는 개인정보의 안전성 확보 조치와 비교된다고 하겠다.

3) 정리

이상 개인정보보호법 및 동법 시행령의 위임을 받은 고시 중 특히 본 연구에서 논하는 주제와 관련된 고시 2건의 내용을 검토하여 보았다. 규범력있는 법규 중 가장 하위 단계에서나마 암호화 관련 규정이 확인되고는 있으나, 이는 이미 구법 시절부터 도입된 것에 불과하다.

암호화의 대상을 구체적으로 특정하였다는 점, 일부 규정의 경우 암호화의 구현방식까지 규정한 점에서 수범자의 예측가능성을 도모하고자 하는 본 연구의 논지에 일부 부합하는 규율이라 하겠으나, 각 고시는 아래와 같은 문제점을 지적할 수 있다.

즉, 첫째, 법규 단계에서 적정한 암호화조치의 기준을 대강이나마 정할 수 있었던 마지막 기회인 고시에서 조차 '안전한 알고리즘', '상용 암호화 소프트웨어' 정도로 암호화조치의무를 규정하여 적정한 암호화조치의 내용을 짐작할 수 있다고 보기는 어려운 점, 둘째, 고유식별정보나 중요한 개인정보 및 가명

정보의 일부 처리 형태에 관하여만 암호화조치의무를 정하고 있어 모든 가명정보 등 개인정보의 처리에 적용되는 규정이 아닌 점, 셋째, 정보통신서비스 제공자가 아닌 개인정보처리자에 대하여는 일정한 경우 암호화조치의무를 면할 여지를 남기고 있는 점, 넷째, 여러 가지 가명처리 방식이 존재하는 상황에서 가명정보와 추가 정보의 적정한 관리 기준을 기존의 고시 내용에 의한다고 규정하였을 뿐, 적정한 「적정한 가명처리」가 무엇인지, 익명정보, 추가 정보의 정의나 판단기준이 제시되지 않는 등 법령 단계에서 공백을 둔 내용에 대하여 규율되지 않은 점, 다섯째, 아래 사례에서 재론하겠으나 비밀번호에 대한 일방향 암호화 그 자체만으로 적정한 암호화조치라고 보기 어려운 경우가 있는 점, 여섯째, 법 제15조 제3항에 따른 이용도 새롭게 도입된 개인정보의 이용형태임에도 이에 대하여는 시행령까지만 추상적인 규율이 이뤄진 결과 고시는 이에 대하여 침묵하고 있는 점 등에서 개인정보 보호에 취약점이 발생할 수 있고, 개인정보처리자 등 정보이용 주체에 대하여 예측가능성을 충분히 보장하지 않고 있다는 측면에서 비판의 소지가 있다.

결국 아래에서 살펴보는 것과 같이 각 개인정보처리자 등이 속한 분야, 업계와 관련되는 주무관청 별로 가이드라인 설정하여 실무적, 기술적인 내용을 사실상 규율하고 있게 되었는데 그 내용과 문제점은 아래에서 항목을 바꾸어 논하도록 하겠다.

다. 각종 가이드라인

개보위나 각 주무관청에서는 위와 같은 법령의 내용에 대하여 세부적인 설명을 하고자 각종 가이드라인, 해설서를 제작, 공표하고 있다.

개보위는 2020. 2. 4. 신법 개정 이후 가명정보 처리 가이드라인을 제작, 공표하였고, 수차례 '개정'이라는 표현을 사용하면서 업데이트하고 있는데, 위 가이드라인에 따르면 보건의료데이터 활용 가이드라인(소관 : 보건복지부), 교육분야 가명·익명정보 처리 가이드라인(소관 : 교육부), 공공분야 가명정보 제공 실무안내서(소관 : 행정안전부), 금융분야 가명·익명처리 안내서(소관 : 금

융위원회)를 가명정보 관련 가이드라인으로 볼 수 있을 것이다.[30]

그러나 이미 수차례 간략히 언급한 것과 같이 가이드라인이나 해설서가 기술적이고 가변적인 내용에 관하여 수범자들에게 일응의 기준을 제시하는 순기능이 있다고 하더라도 그 법적 취급, 효력을 고려할 때 개인정보의 보호, 이용이라는 첨예한 이익이 대립하는 지점에서 충분한 역할을 하고 있는지에 대하여 본 연구의 문제의식이 있다.

그와 같은 견지에서 이하에서는 가이드라인의 법적 성격을 먼저 짚어보고, 일부 가이드라인과 해설서 위주로 내용, 문제점 등을 살펴보겠다.

1) 행정주체가 제작, 공표한 가이드라인의 법적 성격

행정청에서 가이드라인을 제작, 공표하는 경우가 다수 있고, 이러한 가이드라인의 법적 성격은 어떠한지, 법적 취급은 어떻게 하여야 하는지에 대하여 다양한 논의가 있는데, 일반적인 견해는 가이드라인을 행정청이 행정목적 달성을 위하여 기준과 절차를 정한 규범으로 구속력이 없는 행정지도와 유사한 것으로 보는 듯하다.[31]

이와 관련하여 행정절차법 제2조 제3호는 행정지도에 대하여 「행정기관이 그 소관 사무의 범위에서 일정한 행정목적을 실현하기 위하여 특정인에게 일정한 행위를 하거나 하지 아니하도록 지도, 권고, 조언 등을 하는 행정작용」이라고 규정하고 있다.

한편, '가이드라인'의 명칭을 사용하더라도 법률의 근거에 따라 작성되는 경우가 다수 있고[32], 훈령, 예규, 고시의 형태를 취한 예도 다수 있으므로[33], 이

30) 개보위, 위 가이드라인, p.8

31) 이상엽, 망 이용계약 가이드라인 법적 함의와 전망, KISO저널 제37호, p.18, 고학수 등 7명, 앞의 책, p.110, 헌법재판소는 행정청이 행정의 상대방에 방향을 제시하고 자발적인 순응을 유도하려는 가이드라인에 불과하고, 그러한 가이드라인이 행정청의 우월적인 지위에 따라 일방적으로 강제된 것이 아닌 이상 이를 공권력의 행사로 볼 수는 없다고 보았다.[헌법재판소 2021. 11. 25 자 2017헌마1384, 2018헌마90, 145, 391(병합) 결정]

32) 예를 들어 양육비 이행확보 및 지원에 관한 법률 제5조에 따라 여성가족부장관은 '양육비 가이드라인'을 정할 수 있고, 자연재해대책법 제16조의6에 따라 행정안전부장관은 '방재기준 가이드라인'을 정할 수 있다.

33) 대규모유통업 분야에서 납품업자 등의 종업원 파견 및 사용에 관한 가이드라인(공정거래

러한 경우에는 기본적으로 행정규칙으로서 행정청 내부의 의사결정 기준으로서 작용하는 외에 법령의 규정 취지나 행정청의 관행에 따라서는 사실상의 규범력도 인정될 여지가 있을 것이다.[34]

또한 행정규칙의 경우에는 국회법 제98조의2 제1항에 따라 소관 상임위원회에 제출되어야 하고, 상임위원회는 그 중 대통령령, 총리령, 부령의 경우 법률의 취지 또는 내용에 합치되지 아니한다고 판단하는 경우 정부에 통보할 수 있으며 이를 통보받은 중앙행정기관의 장은 처리 계획과 그 결과를 보고하여야 하므로 제한적이나마 국회의 통제도 받게 된다.[35]

따라서 가이드라인은 일반적으로 규범력이 인정되기 어렵고, 행정규칙으로서의 성격도 지니지 않는 경우가 다수 있지만, 그 명칭과 달리 행정규칙의 요건을 갖추고 있는 경우가 있고, 내용에 따라서는 행정의 상대방인 국민이나 법원의 재판에서도 일정한 효력이 있는 경우가 있으며, 사법부, 입법부의 통제를 우회하는 수단으로 사용될 우려가 있는 형태의 행정행위라 하겠다.

위원회예규 제360호), 유비쿼터스 도시기술 가이드라인(국토교통부고시 제2013-390호), 특별건축구역 운영 가이드라인(국토교통부훈령 제1445호) 등

34) 판례는 일반적으로는 행정규칙이 국민이나 법원을 기속하는 대외적 효력을 지니지 않는다고 보고 있으나(대법원 1995. 2. 14. 선고 94누12982호 판결 등), 다만 행정청이 행정규칙에 따라 업무를 반복하여 일정한 행정관행이 이뤄지게 되면 행정청이 자기구속을 받아 이를 위반한 처분에 대하여는 평등의 원칙이나 신뢰보호의 원칙에 위배될 여지는 있다고 보고 있으며(대법원 2009. 12. 24. 선고 2009두7967호 판결 등), 행정처분의 위법성이 아닌 법규의 수범자인 사인의 민, 형사상 책임 유무에 관하여는 고의, 과실 등의 판단에 하나의 기준이 될 수 있음은 명백하다.

35) 국회법 제98조의2(대통령령 등의 제출 등) ① 중앙행정기관의 장은 법률에서 위임한 사항이나 법률을 집행하기 위하여 필요한 사항을 규정한 대통령령·총리령·부령·훈령·예규·고시 등이 제정·개정 또는 폐지되었을 때에는 10일 이내에 이를 국회 소관 상임위원회에 제출하여야 한다. 다만, 대통령령의 경우에는 입법예고를 할 때(입법예고를 생략하는 경우에는 법제처장에게 심사를 요청할 때를 말한다)에도 그 입법예고안을 10일 이내에 제출하여야 한다. (중략) ④ 상임위원회는 제3항에 따른 검토 결과 대통령령 또는 총리령이 법률의 취지 또는 내용에 합치되지 아니한다고 판단되는 경우에는 검토의 경과와 처리 의견 등을 기재한 검토결과보고서를 의장에게 제출하여야 한다. (중략) ⑥ 정부는 제5항에 따라 송부 받은 검토결과에 대한 처리 여부를 검토하고 그 처리결과(송부 받은 검토결과에 따르지 못하는 경우 그 사유를 포함한다)를 국회에 제출하여야 한다. ⑦ 상임위원회는 제3항에 따른 검토 결과 부령이 법률의 취지 또는 내용에 합치되지 아니한다고 판단되는 경우에는 소관 중앙행정기관의 장에게 그 내용을 통보할 수 있다. ⑧ 제7항에 따라 검토내용을 통보받은 중앙행정기관의 장은 통보받은 내용에 대한 처리 계획과 그 결과를 지체 없이 소관 상임위원회에 보고하여야 한다. (후략)

따라서 가이드라인의 내용을 점검할 때에는, 그 형식이 어떠한지, 내용이 상위 법령에 근거하는지, 상위 법령에 상충 되는 것은 없는지. 수범자인 국민에게 지도, 권고, 설명을 넘어서는 사실상의 의무를 부과하는 측면은 없는지 등을 중심으로 검토할 필요가 있다고 본다.

아래에서는 가명처리의 일반 가이드라인인 가명정보 처리 가이드라인과 본 연구에서의 논의에 일부 관련되고, 개인정보이용 요건은 물론 허용 여부 자체에 관하여도 첨예한 대립이 있는 의료 분야에 관한 보건복지부의 보건의료데이터 활용 가이드라인의 내용을 검토하겠다.[36]

2) 가명정보 처리 가이드라인

가) 효력과 적용범위의 문제

위 가이드라인은 '가명정보 처리에 관한 특례', 즉, 법 제3장 제3절에 관한 설명과 구체적 사례를 제공하여 가명정보 처리에 대한 이해를 돕고, 처리 과정에서 발생할 수 있는 개인정보 오·남용을 방지하여 안전한 가명정보 활용 방안을 안내하기 위해 제작되었음을 밝히고 있다.[37]

즉, 위 가이드라인은 신법이 도입한 가명처리 특례에 관한 설명으로 보건복지부가 작성하는 보건의료데이터 가이드라인 등 분야별 가이드라인이 있는 경우에는 보충적으로 적용된다고 하면서도, 법 제15조 제3항에 따른 이용의 경우에는 적용대상이 아니라고 밝히고, 다만 기술적 내용은 참고할 수 있다고 설명한다.[38]

위와 같이 가이드라인은 가명정보의 처리에 대한 이해를 돕고, 안전한 가명정보 활용 방안을 안내한다고 하여 일응 행정지도의 취지와 유사한 목적을 밝

---

36) 개보위 및 보건복지부의 각 가이드라인을 행정지도로 보는 견해가 지배적이라고 소개하는 문헌이 있다. (이석배, '보건의료데이터 활용 가이드라인'의 현행법상 문제점, 대한의료법학회「의료법학」제22권 제4호, p.25) 이하에서는 그 법적 성격을 차치하고 가이드라인의 규율 내용을 위주로 살펴본다.

37) 개보위, 위 가이드라인, p.5

38) 개보위, 위 가이드라인, p.6

히고는 있으나, 한편으로는 가이드라인의 '적용' 범위를 명확히 구분하고, 아래에서 드러나는 것과 같이 대부분의 내용이 적정한 '가명처리'의 예시를 제시하고 그 요건도 다루고 있으므로, 가명처리를 통하여 개인정보를 처리하고자 하는 수범자는 가이드라인을 단순히 참고만 하여도 무방한 것이 아니라 개인정보 처리 시에 가이드라인을 따를 방도 외에 다른 도리가 없을 것이다.[39]

또한, 신법에서 새로이 도입한, 법 제15조 제3항 등에 따른 이용에 대하여는 가이드라인을 단순히 참고할 수 있다고 밝히는 부분도 검토의 여지가 있다. 비록 활용 가능성이 낮다고 하더라도, 법에서 정한 요건이 '정보주체에게 불이익이 발생하는지 여부, 암호화 등 안전성 확보에 필요한 조치를 하였는지 여부'이고 시행령에서는 명시적으로 가명처리 또는 암호화를 하나의 요건으로 들고 있는 점, 가명처리에 의한 이용에 비하여 정보주체, 정보이용주체의 보호 필요성이 높은 점 등에서 위와 같은 경우에도 가이드라인이 일응의 기준을 제시할 필요가 있다 하겠다.

나) 용어 정리

한편, 가이드라인은 개인정보, 가명처리, 추가 정보 등에 대한 용어에 대하여 설명하고 있는데[40], 개인정보, 가명처리, 익명정보의 경우에는 법의 내용을 그대로 옮겨왔다.

위와 같은 용어의 정의 부분에서 주목할 만한 것은 '추가 정보'를 「개인정보의 전부 또는 일부를 대체하는 가명처리 과정에서 생성 또는 사용된 정보로서 특정 개인을 알아보기 위하여 사용·결합될 수 있는 정보(알고리즘, 매핑테이블 정보, 가명처리에 사용된 개인정보 등)로서 가명처리 과정에서 생성·사용된 정보에 한정된다는 점에서 다른 정보와 구분되는 것」으로, '재식별'은 「특정 개인을 알아볼 수 없도록 처리한 가명정보에서 특정 개인을 알아보는 것」으로 각각 정의한 부분이다.

39) 결국 가이드라인도 '개인정보처리자가 법에 따른 규정을 준수한 경우 가이드라인 미준수로 처벌받지 않고, 상황에 따라 유동적으로 처리 가능함'이라고 설명은 하고 있으나(개보위, 위 가이드라인, p.5), 이러한 설명이 수범자에 대하여 예측가능성을 부여한다고 보기는 어렵다.
40) 개보위, 위 가이드라인, p.7

재식별[41]의 경우, 법률에 규정된 문언인 '개인을 알아볼 수'있는 경우를 지칭하는 것으로 아래에서 상론할 비식별조치의 개념에서 논리 필연적으로 도출되는 것인데 적정한 가명처리를 논의함에 있어 중요한 개념이지만, 문언 자체로 일의적이어서 논의할 여지는 없어 보인다.

다만, 추가 정보는 법 제2조 제1호 다목, 제28조의2 제2항, 법 시행령 제29조의5 등에서 '원래의 상태로 복원하기 위한 추가 정보'로 정의되고 있었는데 가이드라인에서 좀 더 나아가 설명하고 있어 검토한다.

일단 추가 정보는 법령(법 제28조의4, 동법 시행령 제29조의5 제1항 등)에 따라 개인정보처리자에 대하여 접근권한의 분리, 분리 보관, 파기 등의 의무가 발생하는 요건이고, 개인정보처리자 등 이용 주체의 입장에서는 중요한 정보이자, 각종 책임의 발생 소지가 있는 것이므로, 위와 같이 그 의미를 구체적으로 정의하고, 포섭 범위를 제한하는 것은 적절한 시도라 하겠다. 예를 들어 가명처리의 한 방식으로 암호화 알고리즘을 상정한다면, 개인정보처리자가 보관하거나, 사용자가 개인정보처리자에게 일시적으로 제공하게 되는 암호화키도 추가 정보의 범주에 들어갈 것이다.

다만, 추가 정보를 「개인정보의 전부 또는 일부를 대체하는 과정에서 생성 또는 사용된 정보」로 덧붙여 설명한 부분은 규범적 판단의 기준으로서 일의적이고, 기술적인 측면에서 세부적이거나 가변적인 사항이 아니며, 추가 정보라는 중요한 법률 요건을 설명한 것이므로 법률의 위임에 따라 법령의 단계에서 규율하는 등으로 규범력을 확보할 필요가 있다고 판단된다.

다) 가명처리와 가명정보의 처리[42]

가이드라인은 가명처리와 가명정보의 처리를 구분하고, 나아가 「가명정보는 가명처리를 수행한 당시의 목적과 처리 환경에 따라 이용하는 것이 원칙」이라고 설명하고 있다.

---

41) 뒤에서 살펴볼 HIPAA의 개인정보규칙(privacy rule)에서는 re-identification의 개념을 명시적으로 규정하고 있다. [45 CFR 164.502(d)(2)(ii)]

42) 이하 개보위, 위 가이드라인, p.10

가명처리와 가명처리된 가명정보의 처리를 구분하는 것은 적절한 조치라고 생각되나, 가명정보의 이용이 가명처리 수행 당시의 목적, 처리 환경에 구속된다고 설명하는 것은 법령에 근거 없이 개인정보의 이용에 제한을 두는 해석이라 생각된다.

그 이유는 첫째, 가이드라인도 자인하고 있듯이 법 제28조의2 제1항 등이 가명정보의 이용 목적 외에 다른 제한을 두고 있지 아니한 점, 둘째, 일단 적정한 가명처리가 된 이상, 법 제28조의7의 특례에 따라 개인정보의 수집 출처 고지의무(법 제20조), 개인정보의 파기의무(법제21조), 영업양도에 따른 개인정보 이전 시 의무(법 제27조), 개인정보 유출 통지(법제34조), 정보주체의 열람, 정정, 삭제, 처리정지요구권(법제35조부터 제37조까지) 등의 규율이 배제되는 점, 셋째, 위와 같이 해석할 경우 법 제15조 제3항 등에 따른 이용과 차이가 없게 되는 점, 넷째, 그 외 법령에서 가명처리 관련 목적, 환경 등에 관한 규율이 이뤄지지 않은 점 등을 고려할 때 입법자가 개인정보보호법을 개정하는 입법을 하였을 때, 가이드라인이 설명하고 있는 내용을 상정하였다고 보기 어렵다.

또한 법률에 없는 근거 없이 가명처리를 제한하면서도 그 목적을 밝히지 아니한 가운데, '목적과 처리 환경'이라는 다분히 추상적인 요건을 두고 있을 뿐이어서 어떠한 효과를 기대할 수 있는지도 알 수 없다.

라) 세부 절차에 대한 설명

그 다음으로 가이드라인은 목적 설정 등 사전준비(1단계) → 위험성 검토(2단계) → 가명처리(3단계) → 적정성 검토(4단계) → 안전한 관리(5단계)의 순으로 단계별 절차도를 그려 가명처리를 도식화하고 있는데[43], 본 연구에서 관심을 가지고 있는 가명처리 관련 내용을 위주로 살펴본다.

① 사전준비 단계

43) 개보위, 위 가이드라인, p.10

가이드라인은 통계작성, 과학적 연구, 공익적 기록보존을 위한 목적의 해석 및 가명정보의 처리, 처리위탁 등을 포함한 계약서 작성의 예시를 들어 설명하여 법 이해에 도움을 주고 있다.[44]

한편, 가이드라인은 「민감정보 또는 고유식별정보도 가명처리의 특례에 따라 가명처리가 가능하나, 주민등록번호는 법령에 근거가 없는 한 가명처리가 허용되지 않는다.」고 설명하여 의료정보와 같은 민감정보도 가명처리의 대상으로 볼 수 있다는 입장인데, 이에 대하여는 적정하게 처리된 가명정보에 대하여는 법의 일부 규정의 적용을 배제하는 법 제28조의7에서 법 제23조(민감정보의 처리 제한)를 적용제외규정에서 포함하지 않고 있으므로 민감정보까지도 정보주체의 동의 없는 가명처리가 가능하다고 해석하는 것은 무리가 있다는 비판이 있다.[45]

위 견해는 보건의료데이터 활용 가이드라인을 비판적으로 고찰하면서 동 가이드라인이 "안전한 가명처리 방법이 현재 개발되지 않은 정보의 경우 가명처리 방법이 개발될 때까지는 (중략) 정보주체 동의하에서 활용이 가능하다."고 설명한 부분[46]을 들어 빅데이터 시대에 안전한 가명처리가 의문이고 기술적 기준이 적절한지도 의문이며 결국 가이드라인이 안전한 가명처리가 없다는 점을 자인한 것이라고 평가한다.[47]

그러나 아래와 같은 이유에서 신법의 입법 취지, 목적 및 법 제2조, 제23조에 규율된 내용을 종합할 때, 일단 민감정보도 가명처리의 대상에 포섭된다고 보아야 할 것이다.

즉, 신법이 적정한 가명처리를 전제로 가명정보의 동의 없는 이용을 전제하고 있는 점, 신법의 입법 목적이 빅데이터의 적극적인 활용, 기술 개발에 있는 점, 법 제23조는 법 제2조 제1호의 개인정보에 대한 처리의 일반적인 상황에 대한 것이고 가명처리는 법 제2조 제2호에서 별개의 처리로 구분되고 있는

---

44) 개보위, 위 가이드라인, p.9-12
45) 이석배, 앞의 논문, p. 21, 계인국, 이성엽, 보건의료 데이터 활용의 법적 쟁점과 과제, 공법연구 제50집 제2호, p.158
46) 보건복지부 등, 보건의료데이터 활용 가이드라인(2021), p. 11
47) 이석배, 앞의 논문, p.20

점, 법 제23조 제1항 제2호에서 「법령에서 민감정보의 처리를 허용하고 있는 경우」를 예외로 들고 있는 점, 민감정보에 대한 특칙이 법 제23조라면 가명정보 관련 규정은 개정으로 추가된 신법으로서 법 제28조의2 이하에서 규정된 점[48], 법 제28조의2에서 가명처리의 목적을 제한할 뿐 처리대상을 제한하지는 않는 점[49] 등에서 후술하는 것과 같이 의료법 등 개별 단행법의 규율이 우선하는 경우를 예외로 하고 나머지의 경우 일응 민감정보도 가명처리의 대상이 된다고 봄이 상당하다.

물론 위와 같이 민감정보에 대한 가명처리 가부에 관하여 견해가 구구한 상황에서 가이드라인이 이에 대한 의견을 제시하는 것만으로는 논란을 해결할 수는 없을 것으로 보이므로 개인정보보호법 단계에서 입법으로 명확한 정리가 필요하다고 판단된다.[50]

② 위험성 검토 단계

위 가이드라인은 가명정보 자체만으로 특정 개인을 알아볼 수 있는지, 가명정보를 처리할 자가 보유하거나 접근, 입수 가능한 정보와의 사용, 사용결합을 통하여 식별할 수 있는지를 토대로 위험성을 검토하여야 한다고 설명한다.[51]

이는 '식별가능성'의 문제로 연결되는 부분인데 식별가능성의 유무에 따라 가명정보로 인정되는지 여부가 결정될 것이다. 즉, 식별가능성이 인정된다면 법에서 정보주체의 동의 없는 개인정보의 처리가 가능한 가명정보로 볼 수 없기 때문에 이러한 경우에는 개인정보의 처리가 위법하게 될 수 있다. 결국 식

48) 이원복, 유전체 연구와 개정 개인정보 보호법의 가명처리 제도, 이화여자대학교 법학논집 제25권 제1호, p.208에서는 민감정보에 대하여 법 제18조 제2항의 개인정보의 목적 외 이용·제공 제한의 예외도 적용될 수 있다고 주장하면서, 법 제18조 제2항이 일반규정이라면 민감정보에 관한 규정이 특칙이라고 보더라도, 가명처리에 관한 규정은 민감정보 등에 대하여도 특칙이 된다고 본다.

49) 김송옥, 앞의 논문, p.399

50) 신법 개정 취지를 고려하여 법체계를 정비할 필요가 있다는 견해가 있다.(계인국, 이성엽, 앞의 논문, p.158) 같은 취지로 법률 개정을 통하여 민감정보는 가명처리하여 활용하더라도 일반 개인정보보다는 더 신중하고 제한적으로 다루어야 한다는 취지의 규정이 필요하다는 견해가 있다.(김현숙, 앞의 논문, p.142)

51) 개보위, 위 가이드라인, p.15

별가능성은 중요한 규범적 판단의 요건인데, 이를 가이드라인의 형식으로 다루는 것은 부적절하다고 본다.

또한 가이드라인의 위와 같은 설명 자체가 불충분하다. 우선 가명정보 자체만으로 특정 개인을 알아볼 수 있는지 여부에 대하여 구체적인 기준 자체를 제시하지 아니하므로 해석과 적용에 논란이 우려된다.

예를 들어 '성기범'이라는 문자열을 가명처리함에 있어 이를 사전에 정한 치환테이블에 의하여 변환하면 '7178124617552)'로 표시할 수 있다고 할 때, 위 문자열을 해시함수 중의 하나인 MD5로 계산한 해시값은 'EF2A2877DBA0E056AE47C07CB98AA8D1'로 표시된다.

일반인이 육안으로 본다면 어느 경우에도 원 데이터가 '성기범'임을 알기는 사실상 불가능하다. 그러나 성명과 가명처리된 결과값의 쌍을 다수 보유하고 있는 공격자에게 적당한 시간을 부여하면 데이터 '71781246175'가 '성기범'을 가명처리한 결과임을 쉽게 파악될 수 있고, 해시함수의 경우에도 뒤에서 살펴보는 것과 같이 암호화 대상인 평문의 길이가 짧거나, 뒤에서 해시함수와 관련하여 살펴볼 '솔트' 등의 추가 정보 처리가 없다거나, 라운딩과 출력비트가 낮은 경우에는 안전성을 담보할 수 없기 때문이다.

결국 식별의 주체, 수단, 난이도에 관한 구체적인 설명이 없다는 사정은 규범력은 물론 수범자의 예측가능성 자체가 현저히 낮아지는 결과를 초래하게 된다. 따라서 어떤 사람의 입장에서 어떠한 대상을 어느 정도의 기술력, 비용 등으로 가명처리하여야 적정한 것인지 추상적인 형태로라도 규율할 필요가 있고 이는 뒤에서 살펴보는 GDPR의 예에서 시사점을 찾을 수 있다고 판단된다. 물론 가이드라인이 가명정보처리자에서 확인할 수 있는 다양한 요소를 식별위험성(가능성)의 판단 기준으로 열기한 것은 긍정적인 평가를 받을 수 있을 것이다.

계속하여 가이드라인은 상정 가능한 데이터 자체의 특성(식별정보, 식별가능정보, 특이정보, 재식별 시 영향도) 및 데이터 처리 환경의 측면(활용 형태, 처리 장소, 처리 방법)으로 구분하여 식별 위험성 요소의 예시를 들어 설명하고

52) 한글자모 중 성명에 사용되는 35개(쌍자음, 겹받침 제외)에 대응하는 숫자를 정하고 풀어쓰기로 한글자모를 분해한 다음 대응하는 숫자를 나열한다고 가정한다.

있는데, 상당한 접근이라고 할 것이나, 마찬가지로 위 분류에서 사용한 개념(식별정보, 식별가능정도, 활용 형태, 처리 방법 등)을 법령 내지 행정규칙의 단계에서 제시하는 것이 입법을 통한 규범력 확보 내지 행정규칙의 형식을 통한 예측가능성의 보장 차원에서 더 나은 방안임은 재차 강조하고자 한다. 위와 같은 사항을 규범화한다고 하더라도 기술의 세부사항을 강제하는 것이 아니므로 기술중립성을 해치는 등의 문제가 야기될 우려는 상대적으로 낮다.

③ 가명처리와 그 기법

가이드라인은 가명정보의 식별 위험성 검토 결과를 기반으로 가명정보의 활용 목적 달성에 필요한 가명처리 방법 및 수준을 정하여 항목별 가명처리 계획을 설정할 필요가 있다고 설명하는데 필수적인 내용이라고 생각된다.[53]

문제는 적정한 가명처리의 판단 기준이 식별가능성에 관한 부분과 마찬가지로 가이드라인에서 조차 설명되지 않고 있다는 것이다. 적정한 가명처리의 판단 기준은 식별위험성(가능성)의 판단 기준에서 살펴본 것과 같이 행정청이나 수범자의 입장에서 적법한 가명정보, 개인정보의 처리, 이용 여부를 판단함에 있어 결정적 요소이기 때문이다.

특히 가이드라인은 가명처리와 가명처리된 결과인 가명정보에 대한 처리를 구별하고 있으므로 각각에 대한 적정한 처리 등 이용 기준을 정립할 필요가 있다.[54]

한편, 가이드라인은 적정한 가명처리의 판단 기준을 제시하지 아니하면서도 부록 1 참고자료란에 개인정보 가명처리 기술 및 예시라는 제목으로 다양한 가명처리 기법을 열거하고 있는데, 크게 개인정보 삭제, 개인정보 일부 또는 전부 대체, 기타 기술로 나누어 설명하고 있다.

이미 수차례 언급한 것과 같이 가명처리, 암호화조치 등의 비식별조치를 하는 이유는 첫째로는 단순히 개인정보를 안전하게 관리하고자 하는 목적이 있

53) 개보위, 위 가이드라인, p.32

54) 가이드라인이 항목을 바꿔서 설명하는 제4단계인 적정성 검토에서도 식별 위험성, 가명처리 계획, 결과에 대한 평가가 이뤄져야 한다고 보고 있으나 그 기준은 구체적으로 제시하지 않고 있다. (개보위, 위 가이드라인, p.34)

을 수 있고, 둘째로는 안전하게 관리되고 있는 정보를 안전하고도 효과적으로 처리하는 등의 방법으로 이를 활용하고자 하는 목적이 있을 수 있다.

경우에 따라서는 첫 번째의 목적만으로도 충분할 수도 있으나, 두 가지 모두를 동시에 추구하여야 할 때도 있을 것이다. 이와 같은 관점에서 가이드라인이 제시하는 가명처리 기법을 살펴본다.

⑴ 「개인정보의 삭제」는 그 정도에 따라서는 익명처리와 동일하게 볼 수 있을 만큼 강력한 처리로서 개인정보의 안전한 관리의 측면에서는 가장 효과적이다. 그러나 개인정보를 삭제하는 방식으로 가명처리를 하는 경우, 개인 식별이 되지 않으므로 정보주체에게 정보처리의 결과를 제공할 수 없고, 데이터 간의 구분이 되지 아니하는 단점이 있다.

그런데 가이드라인은 부분삭제, 행 항목의 삭제를 삭제의 한 유형으로 들고 있으나, 부분삭제의 경우 삭제 후 남은 정보로도 식별가능성이 인정되는 경우가 있을 수 있고, 행 항목을 삭제하는 것은 특정 개인정보를 삭제하는 것에 불과하여 충분한 가명처리로 볼 수 있을지 의문인데 특별한 설명이 없다.

⑵ 「개인정보의 일부 또는 전부의 대체」의 경우, 가장 대표적인 대체 방법인 암호화에 관한 내용을 한국인터넷진흥원(KISA)의 암호이용 활성화 관련 안내서를 참조하도록 전가[55]하고 있어 나머지 부분에 대하여 살펴본다.

마스킹(masking)은 사실상 개인정보의 삭제에 해당하는 것인데 충분한 마스킹이 이뤄진다면 위 개인정보의 삭제와 대동소이한 장단점이 그대로 확인될 것이나, 라운딩, 상하단 코딩의 경우 데이터의 손상이 불가피하다는 면에서 라운딩 등의 정보를 강력하게 하는 경우 효과적 정보처리가 불가능한 경우가 발생할 수 있고, 반면에 이를 약하게 처리한다면 식별가능성이 높아지게 된다.

한편 로컬 일반화, 범주화의 경우 데이터의 이용은 거의 불가능해질 것인데 오히려 성명, 주민등록번호 등 식별자의 삭제로 충분한 효과를 거둘 수 있을 것이므로 실용성이 의문시 된다.

---

55) 개보위, 위 가이드라인, p.78

또한, 잡음 추가, 순열(치환)의 경우, 기본적으로 잡음이 추가 되거나, 순열 등이 이뤄진다면 특정한 데이터의 경우 손상되어 유용성이 낮아지게 되고, 동일한 잡음을 추가하거나, 일정한 규칙에 따라 순열, 치환이 이뤄지게 된다면, 앞서 암호화조치에서 살펴본 것과 같이 몇 가지 평문에 대응하는 암호문의 쌍으로도 공격을 당할 가능성이 매우 높아지므로 유용하지도, 안전하지도 않은 가명처리 방식이라 할 것이다.

결국 가이드라인에서 예시로 들고 있는 가명처리 방안 중, 암호화를 제외한 일반적인 가명처리 방안을 종합할 때, 차라리 정보주체에 대한 식별자(성명, 주민등록번호 등)를 삭제하고 처리하는 방식이 그나마 개인정보 보호 및 이용을 조화롭게 할 수 있을 뿐 나머지 일반적 방식, 라운딩, 일반화, 범주화, 잡음 추가, 순열(치환) 등은 개인정보처리를 통한 활용은 물론 보호에도 그렇게 유용하다고 보기 어렵다.

⑶ 한편, 개인정보처리를 통한 활용의 측면에서 개인정보가 포함된 데이터를 적절히 가공, 생성하여 제공하는 방식을 취한 기술도 예시로 들고 있는데, 이는 수학적, 통계학적 연구를 바탕으로 개진된 것이므로 활용 가치가 충분하다고 판단된다.

우선 가이드라인에서 들고 있는 차분프라이버시(differential privacy)는 아래에서 살펴볼 수리암호 중 블록암호의 형태를 띠고 있는 DES, 스트림암호, 해시함수에 대한 공격방법으로 언급되는 차분 분석(differential cryptanalysis)[56]의 개념을 역으로 개인정보보호에 도입한 것이다.

즉, 개인정보가 담긴 데이터가 총량, 평균으로 제공[57]되었을 때 특정 개인의 정보가 인식, 복원될 수 있으므로 통계적 특성만 유지할 수 있도록 노이즈(잡음)를 추가, 제공하여 데이터 간의 차이가 확률적으로 일정한 크기 이하의 차이를 갖게 함으로써 공격자가 제공된 데이터 간의 차이를 이용하여 개인정보

56) 비트열이 비슷한 평문을 암호화하였을 때 암호문의 비트열(비트列, bit string)에서 나타나는 차이에서 비밀키를 추론해내는 귀납적 방식이다.[김명환, 수리암호학 개론(2019), p.161]

57) 즉, 개인정보를 보유한 신뢰할 수 있는 큐레이터(trusted curator)가 데이터를 분석할 사람(data analyst)으로부터 분석을 요청, 그 결과를 수령하는 사례를 상정할 수 있다. (고학수 등 7명, 앞의 책, p.212)

를 식별하게 되는 가능성을 낮추는 것이다.[58] 이 경우 노이즈[59]를 어떻게 산정하여 추가할 것인지, 이를 공개할 것인지가 쟁점이 된다고 한다.[60]

그 다음으로 가이드라인에서 언급하고 있는 것이 재현데이터(Synthetic Data)인데, 이는 원본 데이터가 아닌 가상의 데이터를 제공하기 위하여 원본 데이터를 분석, 새로운 데이터를 생성하는 기법[61]으로 그 개념대로라면 식별되는 개인정보 자체가 없으므로 이상적인 처리 방식이 될 것인데 폭넓게 활용되고 있는 단계는 아닌 듯하다.[62]

차분 프라이버시나 재현데이터는 확률분포에 따른 노이즈 부여 등 새로운 기법, 통계처리를 이용하는 단계에서의 프라이버시 보호하는 새로운 발상으로 충분한 검토의 가치가 있다고 본다. 다만, 이와 같은 방법은 결국 정보주체의 이익이 되는 각종 개인정보 처리에 활용하기 보다는 대량의 통계를 처리하여 유의미한 귀납적 결론을 도출하는데 사용할 수 있을 것으로 보이는 점, 특히 법이 정한 가명처리, 즉 개인정보를 보관하는 입장에서의 가명처리 의무는 별개로 준수하여야 하므로 위와 같이 데이터를 제공하기 위한 가공 등 처리 시 복호화 등의 문제는 남아 있을 것으로 보이는 점은 짚고 넘어갈 수밖에 없다.

④ 암호화조치와 가이드라인

앞서 간략히 언급한 것과 같이 가이드라인은 암호화조치를 가명처리 중 하나, 즉 개인정보의 일부 또는 전부의 대체 유형으로 분류하고도 개괄적인 암호화 알고리즘을 열거한 채 상세한 장·단점의 설명을 회피하고 있는데[63], 특히 정보주체에 대하여 개인정보의 이용 등 처리 결과를 다시 제공하는 경우에

58) 개보위, 위 가이드라인, p.81
59) 유전정보의 처리의 경우, 식별력이 높은 수치인 단일염기다양성(SNP, single nucleotide ploymorhism)을 비식별화하고자 이에 노이즈를 추가하게 되면 유전정보 분석이 불가능함을 지적하는 견해(이원복, 앞의 논문 p.197)가 있고, 이러한 문제점은 데이터의 특성에 따라 다른 경우에도 발생할 수 있는 것으로서 중요한 지적이다.
60) 고학수 등 7명, 앞의 책, p.218
61) 개보위, 위 가이드라인, p.81
62) 고학수 등 7명, 앞의 책, p.234
63) 개보위, 위 가이드라인, p.67-69

는 복호화가 가능한 암호화조치가 가장 효과적인 수단이므로 그에 대한 상세한 규율이 없는 것은 장차 시정되어야 한다고 본다.

예를 들어 뒤에서 살펴볼 암호화 알고리즘 중 해시함수에서는 해시함수를 적용하는 과정에서 솔트와 같은 추가 정보를 사용하는 것이 안전성 확보에는 필수적인 조치임에도 이를 가이드라인에서는 간략히 예시로만 설명하고 있을 뿐이고, 가명처리와 큰 관련성이 없는 MDC(Message Digest Code), MAC(Message Authentication Code)[64]를 예시로 들고 있는 것은 암호화에 대한 이해가 부족한 결과가 아닌가 생각된다.

또한 수리암호화의 가장 큰 위험성은 비밀키가 내·외부의 공격자들에 의하여 알려지는 것 또는 평문과 암호문의 쌍, 소위 '매칭테이블'과 같은 자료가 누설되는 것 등을 생각해 볼 수 있는데, 이는 법 제28조의2 제2항의 '특정 개인을 알아보기 위하여 사용될 수 있는 정보' 또는 법 제28조의4 제1항의 '원래의 상태로 복원하기 위한 추가 정보'로서 명확히 규율되고 있는 것이다.

문제는 이러한 추가 정보가 어떠한 정보인지 그 판단 기준, 예시가 법령 단계에서 확인되지 않음은 이미 지적하였고, 이러한 정보의 예시야말로 가이드라인에서 '설명'해 줄 수 있는 전형적인 예라 할 것인데, 위 가이드라인에서는 침묵하고 있어 아쉬운 부분이다.

3) 보건의료데이터 활용 가이드라인

위 가이드라인은 개보위와 보건복지부가 공동으로 제작, 공표한 가이드라인으로 신법의 취지를 설명하면서도 「법령에서 구체적으로 정하지 않은 가명처리 등에 있어 보건의료데이터의 특수성을 고려할 필요」에서 제작되었다고 밝히고 있어[65], 상위 법령이 가명처리와 같은 중요하고도 새로운 개념에 대하여

64) 이 함수들은 해시함수의 무결성, 즉 송수신 메시지 간의 차이가 있는지 여부를 판명하는 것에 사용되는 것이다.[이상 https://eprint.iacr.org/2006/294(2023. 1. 3. 확인). Bellare, Mihir; Canetti, Ran; Krawczyk, Hugo, "Message Authentication using Hash Functions — The HMAC Construction", p.1 Appears in RSA Laboratories' CryptoBytes, Vol. 2, No. 1, (Spring 1996)]

65) 보건복지부 등, 보건의료데이터 활용 가이드라인, p.1

구체적으로 정하지 아니한 사정을 확인하면서 특히 보건의료데이터 관련 내용을 설명하는 것에 목적을 두고 있다고 소개한다.

위 가이드라인은 「보건 의료 분야의 개인정보 가명처리 등에 대하여는 동 가이드라인을 우선 적용」[66]한다고 밝히고 있는데, 이는 법 제6조가 「개인정보 보호에 관하여는 다른 법률에 특별한 규정이 있는 경우」 다른 법률이 우선한다는 원칙을 밝히고 있는 것과 상통한다.

보건의료 분야에서 다뤄지는 개인정보의 상당 부분은 법 제23조의 '민감정보' 등 특별히 보호 필요성이 있는 분야임과 동시에 개인정보의 이용을 통한 기술 개발, 연구의 여지가 넓다는 측면[67]에서 위 가이드라인의 내용을 전체 법체계와 비추어 고찰할 필요가 있다고 생각한다.

가) 적용 대상정보의 범위

가이드라인은 법 제23조에 따른 민감정보 중 '건강'에 관한 정보를 대상으로 삼으면서 그 예시로, 「의료법」상 진료기록부 및 전자의무기록, 그밖에 병원 내에서 생산되어 진료내역을 표시하고 있거나 쉽게 추정할 수 있는 기록(영수증 등), 건강보험공단, 기타 민간보험사 등에서 수집한 보험청구용 자료, 가입설계에 사용된 건강·질병·상해 등 관련 자료 및 그 부속자료, 건강검진결과 정보, 의사의 진단, 의료기기의 계측, 알고리즘 등의 추정을 통해 파악·추정한 건강상태 정보, 건강상태, 건강습관 여부·정도를 측정하기 위해 기기를 통해 수집한 정보 및 일반적으로는 건강정보로 보기 어렵지만, 질환의 진단·치료·예방·관리 등을 위해 사용되는 정보 등을 열거하고 있다.[68]

이상과 같이 열거한 정보는 그 작성, 수집, 처리 주체를 막론하고 건강과 관련된 정보 일체를 규율하고자 노력한 것으로 보이는데, 민감정보에 해당하는

66) 보건복지부 등, 위 가이드라인, p.3

67) 이석배, 위 논문, p.6에서는 전 국민이 건강보험 가입이 의무화되어 있고, 보건의료정보가 국민건강보험공단, 국민건강보험심사평가원 등 공공기관이 보유, 관리하고 있어 보험료, 건강검진데이터, 진료정보, 처방내역 등이 체계적으로 보관되어 있으므로 빅데이터로서의 가치와 잠재력이 있음을 확인하면서도, 안정성의 측면에서 위험성이 있음을 강조하고 있는데 적절한 지적이라 생각한다.

68) 보건복지부 등, 위 가이드라인, p.9

것이니만큼 최대한 특별한 관리를 해 보고자 하는 시도로 보인다.[69]

한편, 정보주체의 인권 및 사생활 보호에 중대한 피해를 야기할 수 있는 정보에 대하여는 본인 동의를 받아 활용하는 것을 원칙으로 삼겠다고 하는데[70], 일단 개인정보보호법에는 이와 같은 내용이 명시적으로 규율되지 아니하여 다소 의문이나, 아래에서 보는 것과 같이 의료법과의 관계에 비추어 보면 일응 타당한 면이 있다고 본다.

다만, 인권 및 사생활 보호에 중대한 피해를 야기할 수 있는 정보라는 추상적 규율은 법률 단계에서나 가능한 것이지 이와 같은 설명으로는 수범자들에게 아무런 '가이드라인'을 제공할 수 없는 가이드라인에 불과하다는 것은 밝혀두고자 한다.

나) 가명처리 원칙

가이드라인은 안전한 가명처리 방법이 있을 경우 개인정보를 가명정보로 변환하여 활용 가능하다고 보면서도, 「안전한 가명처리 방법이 현재 개발되지 않은 정보」에 대하여는 그러한 가명처리 방법이 개발될 때까지는 가명처리 가능여부를 판단할 수 없으며 이러한 정보는 정보주체 동의하에서만 활용이 가능하다고 보고 있다.[71]

그러나 '안전한 가명처리 방법이 있을 경우', '안전한 가명처리 방법이 현재 개발되지 않은 정보의 경우'라는 추상적인 요건을 두어 사실상 가명처리를 제한하고 있는 것은 일단 그 자체로 구체적인 기준이라고 볼 수 없는 점, 앞서 살펴본 것과 같이 가이드라인은 법령에 근거를 두지 않은 등으로 규범력이 희박한 점, '안전한 가명처리 방법'의 판단 주체, 기준을 두지 않은 상태에서 이를 요구하는 것은 사실상 가명처리를 허용하지 아니하는 결과를 야기할 수도 있는 점[72] 등을 고려하면 개인정보의 이용을 효과적으로 규율하지 못하고 있

69) 보건의료데이터를 보건의료기본법상 보건의료정보의 정의 규정을 보건의료데이터로 보는 견해(정영진, 앞의 논문, p.207)도 있는데 위 가이드라인도 유사한 입장으로 보인다.
70) 보건복지부 등, 위 가이드라인, p.10
71) 보건복지부 등, 위 가이드라인, p.9
72) 예를 들어 "빅데이터 시대에 안전한 가명처리'가 가능한지 의문이고, 기술적 기준이 적절

음은 물론 법에서 허용하고 있는 가명처리를 사실상 임의로 제한하고 있다는 비판[73]이 가능하다.

다) 가명처리 방법

① 식별자 : 삭제 또는 일련번호로 대체할 것을 요구

우선 가이드라인은 식별자(identifier)라는 개념을 명시하면서 법령상 고유식별번호, 보험번호, ID 등과 같은 식별자는 삭제 또는 일련번호로 대체할 것을 요구[74]하는데, 식별자를 양방향 암호화 하거나 해시함수를 이용하여 암호화한 경우에도 식별자로 간주하면서 가명처리를 하더라도 이를 삭제하여야 한다고 설명[75]한 부분이다.

이는 기본적으로 적정한 암호화조치를 가명처리 방식으로 보고 있는 우리 법령의 문언을 무시한 법령에 근거가 없는 해설에 불과하고, 암호화조치에 대한 이해가 없는 것에 기인한 것으로 보인다.

식별자라는 개념이 법령에 정해진 것이 아니고, 암호화한 정보가 식별자라고 보는 것도 독단적인 견해이다. 적정한 암호화가 이뤄진 경우까지 이를 식별자라고 볼 근거가 없고, 가이드라인이 제시하고 있는 대안인 「일련번호로의 대체」를 가장 용이하고도 안전하게 이행할 수 있는 방법이 암호화이기도 하다기 때문이다.

즉, 암호화키 등 추가 정보 등에 대한 적법한 권한이 없는 자가 합리적 시간, 비용 및 현재 사용가능한 기술로도 복호화할 수 없고, 그러한 추가 정보가 안전하게 보관된 경우에까지 삭제를 요구하는 것은 개인정보의 이용에 대한 합리적 근거가 없는 부당한 제한이다.

한 것인지에 대하여도 의문이다."고 지적하는 견해도 있다. (이석배, 앞의 논문, p.20) 그러나 '안전한 가명처리'라는 표현으로 적정한 가명처리의 인정 가능성 자체를 배제하겠다면 신법 개정 취지를 몰각하는 것으로서 받아들이기 어렵다.

73) 특히 유전체 정보에 대한 가명처리 가능여부를 '적절한 가명처리 방법이 개발될 때까지 유보'한다는 것에 의문을 제기하는 견해로는 이원복, 앞의 논문, p.216 참조

74) 보건복지부 등, 위 가이드라인, p.12

75) 보건복지부 등, 위 가이드라인, p.13

덧붙여 식별자를 삭제한다면 나머지 데이터만으로는 정보주체가 식별될 가능성이 없다면, 그 때부터는 익명정보로서 개인정보보호법의 적용대상에서 아예 제외될 수도 있을 것이기 때문에 이러한 설명이 필요한지도 의문이다.

② 속성값에 대한 상세한 처리방법 규율

다만, 위 가이드라인에서는 각종 속성값, 즉 데이터의 종류에 따라 상세한 처리 예시를 들고 있는데, 이는 평가받아야 할 부분이다.[76]

즉, 가이드라인은 측정수치의 정보, 의료인의 관찰, 입력 정보, 알고리즘 정보에 대하여는 일정한 가명처리가 되었다는 전제 하에 별도의 조치가 필요하지 않고, 체내·외를 촬영한 영상정보는 별도의 조치를 상세히 요구하고 있으며, 음성, 지문, 유전자 등에 대하여는 그 특성에 따라 가명처리를 원칙적으로 금하고 있다.[77]

아래에서 살펴보는 것과 같이 개인정보보호법에 우선하여 의료법이 적용되는 이상, 개인정보의 유형에 따라 정보주체의 동의 없는 가명처리를 금지할 여지가 있다고 볼 수 있고, 특히 민감정보에 해당하고 다양한 유형의 정보와 기술이 교차하는 의료 관련 데이터에 대하여 유형별, 기술별로 일응의 기준을 제시한 것은 적절한 조치라고 생각되며, 이는 개보위의 가명정보 처리 가이드라인도 한번쯤 살펴보아야 할 예시가 아닌가 생각된다.

라) 의료법과 개인정보보호법과의 관계 설정

이미 앞에서 몇 차례 언급한 것과 같이 법 제6조는 개인정보보호에 관하여는 다른 법률에 특별한 규정이 있는 경우를 제외하고는 이 법이 정하는 바에 따른다고 규율하여 타법에 개인정보보호에 관한 규정이 있다면 당해 법률이 우선함을 선언하였다.

---

76) 같은 취지의 견해로는 이원복, 앞의 논문, p.216

77) 이상 보건복지부 등, 위 가이드라인 p.14-16 다만, 유전 정보에 대하여 가명처리를 자의적으로 금지하고 있다는 비판도 있다.(이원복, 앞의 논문, p.217)

이에 대하여 의료법 제21조 제2항은 「의료인, 의료기관의 장 및 의료기관 종사자는 환자가 아닌 다른 사람에게 환자에 관한 기록을 열람하게 하거나 그 사본을 내주는 등 내용을 확인할 수 있게 하여서는 아니 된다.」다고 규정하여 원칙적으로 환자에 관한 기록을 타인에게 제공할 수 없도록 하면서, 같은 조 제3항에서 예외를 열거하고 있으나 환자의 일정 범위 내의 친족이나 대리인의 요청, 각종 의료보험, 보상제도 관련 이용, 범죄수사, 재판, 병역 등 극히 제한적으로 예외를 허용하고 있을 뿐이다. 같은 맥락에서 같은 법 제21조의2는 진료기록의 경우 다른 의료기관, 의료인의 요청 및 환자의 동의를 요건으로 제공을 허용하고 있다.

결국 의료법상 '환자에 관한 기록', '진료기록'은 정보주체인 환자 등의 동의가 없는 한 제공은 물론 열람이나 내용 확인도 불가능하다. 덧붙여 같은 법 제18조 제3항, 제19조 제1항, 제23조 등에서도 전자처방전, 진료정보나 의료인이 자신의 업무 수행상 취급, 지득한 개인 정보를 누설하는 등의 행위를 금하고 있다. 따라서 위와 같은 의료법의 규정에 포섭되는 개인정보에 대하여는 정보주체의 동의가 없는 한 개인정보보보호법상 허용된 각종 정보 처리에 관한 규정에 따라 정보 처리를 할 수는 없다고 봄이 상당하다.

이 문제에 대하여 가이드라인은 '의료기관이 보유 중인 환자에 관한 기록을 제3자에게 제공하는 경우'에만 위 의료법 규정을 적용한다는 해석[78]하여 상당한 검토의 여지를 남기고 있다. 그에 따라 가이드라인은 가명처리를 하여 식별가능성이 없는 진료기록이나 의료기관이 아닌 자가 보유하는 진료기록은 개인정보보호법에서 정한 일반 원칙에 따라 이용할 수 있다고 보고 있다.[79]

본 연구에서와 같이 신법 시행 이후의 법 해석, 적용에 있어 개인정보의 적정한 이용이 하나의 중요한 입법 목적이라고 보는 입장에서도 위와 같은 가이드라인의 입장은 논란의 여지가 있다.

① 의료인, 의료기관이 보유 중이나 가명처리한 진료기록

78) 보건복지부 등, 위 가이드라인 p.34
79) 보건복지부 등, 위 가이드라인 p.34

우선 가명처리하여 식별가능성이 없는 진료기록이라 하더라도 의료기관이 보유 중인 기록이라면 의료법 규정이 적용되어야 하므로 위와 같은 가이드라인의 설명은 큰 오해를 낳을 수 있다.

첫째, 기본적으로 가이드라인은 규범력이 미약하고, 위 가이드라인은 개인정보보호법은 물론 의료법에서 위임을 받지 않은 것이다. 그럼에도 개인정보보호법 제6조에 따라 의료 관련 개인정보보호에 있어서는 우선 적용되는 의료법이 가명정보, 가명처리에 대하여 규정을 하지 않았음에도 위와 같이 가명처리된 경우 의료법의 적용이 배제된다고 단정한 것은 무리가 있다고 판단한다.

둘째, 가이드라인은 식별자에 대하여는 가명처리도 소용없다는 독창적인 견해를 내세우고 있었는데, 이 부분에서는 가명처리를 하면 환자식별력이 없다는 취지로 반대의 견해를 제시하고 있으므로 결국 가명처리에 대하여 일관된 인식을 가지고 있다고 보기 어렵다.

결국 이 부분에 대하여는 상당히 아쉬운 결론이지만, 의료법상 명시적 근거가 없는 한 의료법이 적용되는 환자에 관한 정보, 진료기록에 대하여는 적정한 가명처리가 보장된다고 하더라도 이를 동의 없이 이용하는 것은 의료법의 규정에 반한다고 봄이 상당하고, 이는 아래에서 정리하는 것과 같이 보완 입법으로 바로잡을 문제이다.

② 의료인, 의료기관이 보유하지 않는 정보

한편, 의료법이 의료인과 의료기관이 의료행위 등을 수행하면서 생성되는 정보에 대하여 1차적으로 규율하는 것인 점, 의료법에서 건강, 의료에 관한 정보의 처리, 이용에 대하여 구체적으로 규율하지 아니하는 점, 의료법은 단지 의료인, 의료기관에 대하여 그들이 직접 생성, 관리하는 정보에 대하여 환자의 동의 없이 정보를 처리하지 않을 의무를 부과하고 있는 점 등에서 의료법의 수범자가 아닌 공공기관[80] 기타 법령에 근거하여 의료 관련 정보를 처리하는

80) 정영진, 앞의 논문, p.208 표1에 따르면 국민건강보험공단, 건강보험심사평가원, 질병관리청, 국립암센터의 공공영역, 제약회사의 의약품 정보, 연구소의 실험정보 등을 보건의료데이터의 예로 들고 있는데, 이들에 대하여 의료법이 직접 적용되지는 않을 것이다.

주체[81]에 대하여는 의료법의 제 규정이 당연히 우선 적용된다고 할 수는 없다.

이에 대하여 가명정보 역시 개인정보라는 점에서 가명처리된 진료기록이라면 모두 의료법이 우선 적용되어야 한다는 견해가 제기되나[82], 의료법에서는 가명정보나 개인정보에 대한 규율이 없음은 살펴본 것과 같고, 위와 같이 의료법상 각종 금지 규정은 의료인, 의료기관이 정보주체인 환자의 동의 없이 정보를 제공하는 것을 금지하거나, 의료인, 의료기관이 보관하고 있는 정보를 부정한 수단으로 유출하는 행위를 금하고 있을 뿐이어서 적정한 가명처리가 된 건강에 관한 정보 중 의료법의 적용 범위에서 벗어나는 정보는 개인정보보호법상의 가명처리에 관한 특례 규정에 따라 이용할 수 있다고 본다.[83]

라. 검토

---

81) 예를 들어 공공데이터의 제공 및 이용 활성화에 관한 법률에서는 공공데이터를 「공공기관이 법령 등에서 정하는 목적을 위하여 생성 또는 취득하여 관리하고 있는 광(光) 또는 전자적 방식으로 처리된 자료 또는 정보로서, 가. 전자정부법 제2조 제6호에 따른 행정정보, 지능정보화 기본법 제2조 제1호에 따른 정보 중 공공기관이 생산한 정보, 공공기록물 관리에 관한 법률 제20조제1항에 따른 전자기록물 중 대통령령으로 정하는 전자기록물로 열거 하고 있다.(동법 제2조 제2호) 이에 따라 건강보험공단이나 건강보험심사평가원이 공공데이터의 형식으로 보유하고 있는 정보를 제공할 수는 있다고 본다. 서울특별시에서 운영하는 공공자전거 따릉이의 이용 결과가 '서울 열린데이터 광장'이라는 웹페이지에서 공공데이터의 형식으로 제공되는 사안에 대하여 공개된 정보만으로 '탄소량 계산식'을 도출한 후, 이를 이용하면 이용자들이 공개한 이용내역을 토대로 몸무게 정보를 추산할 수 있다는 흥미로운 결론을 도출한 연구가 있었다.(천지영, 노건태, 데이터 3법 시대의 익명화된 데이터 활용에 대한 제언, 정보보호학회논문지 제30권 제3호, p.507-509) 위 연구에는 공공데이터의 제공으로 야기될 수 있는 사례를 들어 프라이버시 침해 문제가 제기될 것을 경계하고 있다.(천지영, 노건태, 앞의 논문 p.501, 다만, 정보주체의 인적사항 등 식별자가 완전히 삭제되었다면 극단적인 과체중 또는 저체중이 아닌 한 몸무게만으로 식별가능성이 있는 경우는 없을 것으로 보인다.)

82) 이석배, 앞의 논문, p.22

83) 생명윤리 및 안전에 관한 법률은 인간대상 연구의 경우 개인정보의 보호 규정을 특별히 두고 있는데, 일단 의료법과 개인정보보호법의 관계에서와 같이 위 법이 규율하고 있지 않은 부분에 한하여 개인정보보호법이 적용될 수 있을 것이라는 견해도 있다.(이원복, 앞의 논문, p.217) 계속하여 이 견해에서는 유전체의 정보의 예를 들면서, 그러한 정보를 수집한 목적이 연구의 목적이라면 위 법이 적용될 것이나, 희귀병의 진단을 위하여 수집되었다면 개인정보보호법이 적용될 수도 있다고 본다.(이원복, 앞의 논문, p.218) 개인정보보호법과 의료법과의 관계에서도 참고할 수 있는 논증이라고 생각된다.

1) 법규명령, 행정규칙 등의 단계에서 규율 미진

이상 살펴본 것과 같이 신법은 가명정보, 가명처리, 정보주체의 동의 없는 개인정보의 이용 등의 개념을 야심차게 도입하여 적극적인 개인정보 활용을 뒷받침하고자 개정된 것으로 그 개정경위, 취지에는 대체로 공감한다.

그러나 신법은 정보주체의 동의를 여전히 개인정보처리의 원칙적인 전제로 고수할 수밖에 없는 상황에서 위와 같이 중대한 예외를 도입한 것인데, 법령에서 각종 정의, 기준에 관한 규정을 추상적으로 마련하거나 심지어 상세한 규율을 회피하였음에도 하위 법령에 위임조차 하지 않았다. 이와 같은 문제점으로 인하여 신법의 실효성이 크게 저감되었다고 평가된다.

또한, 이와 같은 문제점은 신법의 위임을 받은 시행령, 고시 등 하위규범의 단계에서라도 보완할 수 있었을 것인데, 이미 살펴본 것과 같이 신법에 따른 시행령, 고시의 내용을 종합하여도 법률에서 위임한 내용을 그대로 반복하거나, 세부사항을 규정하지 아니하였거나, 구법 하의 내용을 그대로 유지하고 있어 모법이 개정된 상황을 제대로 반영하고 있다고 보기는 어렵다.

요컨대 최소한 식별가능성, 적정한 가명처리, 익명처리의 판단 기준, 추가정보의 개념 등 법률 단계에서 이미 규정되어 있고, 각종 개인정보의 처리가 적법한지 여부에 대한 규범적 판단의 요건(예를 들어 식별가능성이나 가명처리의 적정성의 판단 기준에 있어서는 식별의 주체, 수단, 난이도 등)에 대하여는 시행령이나 고시의 단계에서 예시적 열거[84] 등과 함께 상세한 규율을 통하여 입법을 보충할 필요가 있다. 그에 대한 시사점은 아래 외국 법제에서 쉽게 찾을 수 있다.

덧붙여 민감정보의 가명처리 가부, 의료법과의 관계 등과 같이 이미 학계에서 치열한 논의가 이뤄지고 있는 부분들은 입법의 불비 내지 단행법간의 충돌이 발생한 경우이다. 이러한 문제에 대하여 법적근거가 없고 규범력이 희박한 가이드라인으로 설명을 시도하는 것은 큰 효과를 기대할 수 없다.

---

84) 마찬가지로 가명처리 기법의 내용이나 장단점을 일일이 열거하지 않더라도 '개인정보 삭제, 개인정보의 일부 전부를 대체하는 삭제, 마스킹, 라운딩, 암호화 알고리즘 등'으로 규율하는 것은 기술중립성의 문제가 거론되지 않을 것이다.

결국 위와 같은 문제는 각종 입법조치로서 정리할 수밖에 없다는 결론에 이르렀다. 특히 의료법의 적용 범위에 명확히 포섭되는 정보에 대하여는 현행 의료법의 해석 상 가명처리 등의 조치를 하더라도 정보주체인 환자의 동의 없이는 이를 이용할 수 없다고 새김이 상당하므로, 만약 이 부분에까지 개인정보보호법의 일반원칙을 관철하고자 한다면 사회적 공감대를 형성한 후에 의료법의 개정 등 입법 조치가 반드시 필요하고, 의료인, 의료기관이 보관, 보유하지 않고 있는 '건강에 관한 정보'에 대하여도 개인정보보호법을 직접 적용할 수 있음을 확인하거나, 개인정보보호법상 각종 조치를 준용할 수 있음을 명확히 하는 입법 조치가 있어야 불필요한 논란을 방지할 수 있을 것이다.

2) 가이드라인의 한계와 보완점

가) 규범력의 확보 필요성

신법에서도 여전히 개인정보의 위법한 이용에 대하여는 형사처벌, 과태료 등 각종 제재가 가능하고, 민사책임이 발생할 수 있음에도, 개보위, 보건복지부 등 관계 기관은 구법 시행 당시와 마찬가지로 가이드라인 등을 통하여 신법을 소개하고 있다.

수범자 입장에서는 위와 같이 다양한 법적 위험이 상존하고 있으므로, '가이드라인'을 단순한 권장, 설명이 아닌 일응의 적법성 판단은 기준으로 여길 것이어서 가이드라인을 통상의 행정지도로 여길 수만은 없을 것인데, 상위 법령에 근거를 찾아볼 수 없고, 행정규칙의 형식을 취하지 아니한 가이드라인으로 신법을 설명하고 있는 현 실태는 비판받을 소지가 크다.[85]

이에 대하여 개인정보보호법의 특성상 가이드라인이 불가피하다는 입장에서는 가명처리 등에 관한 요소들이 법 규정에 담아내기에는 상당히 전문적, 기

85) 이에 대하여 법에서 개보위에 법령의 해석 및 운용 권한(법 제7조의9 제5호)을 부여하고 있으므로 법적 근거는 있다고 평가하는 견해(이원복, 앞의 논문 p.212)가 있기는 하나 「개인정보보호위원회는 개인정보보호법 중 적법한 가명처리에 관한 세부사항을 가이드라인으로 정할 수 있다.」는 것과 같은 법령상의 근거를 둔 경우와 비교할 때, 규범력, 구속력에서 큰 차이가 있을 수밖에 없다.

술적일 뿐만 아니라, 관련 기술이 계속 진화하고 있으므로 세부적인 사항에 관하여는 수정이 용이한 형식을 따를 필요가 있고, 오히려 정부가 구속력 없는 가이드라인을 활용함으로써 관련 분야의 특수성을 반영하는 등으로 기술 발전을 저해하지 않을 수 있는 등 법령의 경직성을 극복할 수 있다고 역설한다.[86]

그러나 위와 같은 견해가 현재의 가이드라인 자체로도 충분하다는 주장이라면 찬동할 수 없다.

개인정보보호법이 정보주체와 정보이용주체의 기본권이 충돌할 수 있는 지점을 규율하고 있는 점, 우리의 법현실상 개인정보의 처리를 통한 활용에 대한 불안한 시선이 상존하는 점, 그에 따라 개인정보보호법상 의무위반은 민형사, 행정상 제재로 이어질 수밖에 없는 점 등에서 여타 기술적 분야의 규율 방식과는 다를 수밖에 없다.[87]

더욱이 가이드라인을 발간, 공표하면서 법령의 위임 없이, 행정규칙의 형식을 취하지도 아니한 것은 사회적, 법적 논의를 우회하고 국회의 통제를 피하면서도 법령의 실효성을 확보하려는 시도로 비판받을 수 있고, 수범자인 국민의 기본권을 보장하고 그 행사를 촉진할 국가의 의무를 회피하였다는 평가를 받을 여지가 크다.

이미 살펴본 각 가이드라인에서는 일부 기술적인 분야에서 적절한 설명을 개진하고는 있으나, 법 해석, 적용에 있어 결정적으로 기능하는 부분에 있어서는 언급을 회피하거나, 특정한 방식을 권장한다고만 설명하는 것에 그치고 있다. 이러한 가이드라인으로는 수범자에 대하여 규범력은 물론 어떠한 예측가능성도 보장하지 않고, 오히려 악의적인 수범자에게는 탈법행위를 저지를 구실을 제공할 우려가 있다.[88]

---

86) 고학수 등 7명, 앞의 책, p.112

87) 오히려 훈령, 예규 등의 지위를 가진 가이드라인조차, 기록물 관리, 각종 공공 건축의 디자인, 계약 상 주의사항 등 상당히 기술적인 내용을 설명하고 있다는 점을 상기해 보자.

88) 이석배, 앞의 논문, p.26에서는 보건의료데이터 활용 가이드라인이 “가명정보를 최초 제공받을 당시 원 개인정보처리자에게 밝힌 목적 외의 목적으로 처리할 경우 원 개인정보처리자에게 고지할 것을 권장”, “데이터 분석 대행 또는 협력 연구를 통해 익명정보 반출만으로도 목적을 달성할 수 있을 경우 원 개인정보처리자가 이러한 작업을 수행할 것을 권장”하는 등의 표현을 대표적인 예로 들고 있다.

물론 모든 세부적, 기술적 사항을 고시 이상의 행정규칙에 담는 것은 기술중립성의 원칙에도 반하고, 기술의 발전을 저해할 것이며 과도한 규제로 기능하여 정보이용 주체의 재산권 등 행사의 자유를 제약할 것이라는 우려에는 공감하고 있으므로, 앞서 보건의료데이터 활용 가이드라인에서 확인한 개인정보의 형태별 가명처리 필요성, 가능성을 설명한 부분과 같은 것이 적정한 가이드라인의 전형이라고 평가할 수 있다.

가이드라인의 효용성을 주장하는 견해에서 특히 '구속력 없는 가이드라인이 (수범자들로 하여금) 자율성을 보장'한다는 표현은 오히려 가이드라인이 지닌 맹점을 드러낸다고 생각한다. 가이드라인이 구속력을 가진다면 그 1차적 효력은 규제 당국에 대한 구속력으로 당국이 일관된 규제 행정을 통하여 행정 상대방에게 예측가능성을 부여하게 되는 장점도 있기 때문이다.

나아가 법령의 위임 내지 근거를 지닌 가이드라인은 행정의 상대방은 물론 사법기관에 대하여 일정 수준의 규범력을 가지게 될 것이어서 수범자들이 적정한 가명처리 등과 같이 각종 기술 개발에 심혈을 기울이는 분야에 대하여는 각자가 개발, 이용한 기술이 적법, 적정한지 판단기준으로 주시하게 될 것이고, 결국 새로운 기술 연구, 응용에 있어 안전판으로 삼게 될 것이다.

요컨대 최소한 가명처리에 관한 가이드라인에 대하여는 우선 그 존재에 대한 법령의 위임 내지 근거(예 : 가명처리의 상세한 내용에 관하여는 개인정보보호위원장이 가이드라인을 정할 수 있다.)를 마련하고, 그 내용의 측면에서도 기술적인 부분 이외의 원칙적 내용은 법령의 단위에서, 기술적인 내용은 법령과 상충하는 부분이 없는지를 검토한 후 가이드라인에서 각각 상세히 규율할 필요가 있다. 물론, 이와 같은 내용은 정기적으로 검토의 주기를 정하여 기술의 발전, 이용 상황을 적시에 반영하는 노력도 함께 진행되어야 할 것이다.

나) 상위 법령과의 상충 해소 필요

이상 정리한 것과 같이 가이드라인의 태생적 한계를 인정하더라도, 두 가이드라인에는 아래와 같이 보완할 부분이 다수 확인된다.

① 법령과의 상충, 법령에서 예정하고 있지 않은 의무 부과

가명정보 처리 가이드라인에서는 가명처리 시의 목적, 처리 환경에 따라 가명정보를 이용하는 것이 원칙이라고 설명하나 이는 법령 상 근거가 없고 그 목적도 불분명하므로 재검토가 필요하고, 민감정보에 대한 가명처리 여부, 의료법 등 단행 법률과의 관계는 가이드라인에서 다룰 성격의 문제가 아니므로 위에서 지적한 것과 같이 법령 단계에서 명확히 규율할 필요가 있다.

보건의료데이터 처리 가이드라인에 대하여는 '안전한 가명처리 방법, 안전한 가명처리 방법이 개발되지 않은 경우', '정보주체의 인권 및 사생활 보호에 중대한 피해를 야기할 수 있는 정보에 대하여는 정보주체의 동의를 원칙으로 함' 등의 표현과 같이 법령에 근거를 두지 아니한 채 정보이용 주체에 대하여 사실상의 의무를 부과하고 규범적 판단에 해당하는 부분을 임의로 재단하고 있는데, 위와 같은 규율이 필요하다면 반드시 규범력이 있는 법령 단계에서 정함이 상당하다.[89]

또한, 식별자는 가명처리하더라도 반드시 삭제하여야 한다는 설명은 가명처리에 대한 개인정보보호법의 규정과 상충함이 명백하고 가명처리 기법의 본질을 오해한 설명이므로 수정이 필요하다.

② 가명정보 처리 가이드라인의 내용 보강

법 제15조 제3항 등에 따른 이용의 경우, 시행령이 적정한 조치의 요건으로 가명처리 또는 암호화조치를 요구하고 있으므로 관련 내용 또한 가이드라인의 적용 범위에 포함함이 상당하고, 암호화 알고리즘을 포함한 다양한 가명처리 방식의 장단점, 추가 정보의 예시 및 관리 방안, 과학적 연구의 예시[90]에 대하여는 가이드라인답게 기술적인 내용을 보강하여 상세히 설명할 필요가 있음을 밝혀 둔다.

---

89) 같은 취지로 보건복지부의 위 가이드라인에 법률 또는 적어도 시행령에 규정함이 상당한 내용이 포함되어 있음을 지적한 견해가 있다.(조성훈, 앞의 논문 p.249)

90) 이원복, 앞의 논문, p.216

Ⅲ. 외국의 가명처리 등

이상 2020. 2. 4. 개정된 신법, 동법 시행령, 개보위의 고시 및 가이드라인의 관련 내용을 검토하여 본 결과, 새로이 도입된 개념의 정의, 요건에 대하여 법령의 단계에서 법적 규율이 불충분한 부분을 발견하였고, 특히 규범력이 인정되기 어렵고, 법적 근거가 부족한 가이드라인의 형식으로 가명처리에 관한 중요한 개념, 요건을 설명하면서도 그 중 일부 내용은 상위 법령과 불합치하거나, 부실하게 설명된 부분을 지적할 수 있었다.

이러한 문제점은 개인정보의 적정한 이용은 물론 개인정보의 안전한 보호에 걸림돌이 될 우려가 있고, 수범자에 대하여는 예측 가능성을 저해하고 있으며 행정, 사법기관에 대하여는 적정한 법의 해석, 적용, 집행을 어렵게 할 소지가 있다.

여기에서는 유럽의 GDPR, 미연방의 HIPAA 등을 위주로 본 연구에서 관계되는 부분을 정리하여 위와 같이 지적된 문제점을 보완할 때 필요한 착안점을 모색해 보고자 한다.

1. 유럽 : GDPR의 검토

가. 개요

GDPR은 유럽연합에서 제정한 개인정보보호의 일반법으로서 가입국 내에서는 별도의 입법 없이 법률로 인정되는 것인데, 기존의 유럽연합 정보보호지침[91]이 각 회원국이 국내에서 별도 입법을 해야 하였던 지침(directive)이었다는 점에서 명확히 구분된다.[92] GDPR은 2016. 4. 14. 제정되어 2018. 5. 25.부터 시행 중이다.

---

91) 흔히 'Data Protection Directive'로 지칭되고 있으나, 공식명칭은 'Directive 95/46/EC'이다.
92) 유럽연합 전역에 통일된 개인정보보호법제를 수립, 시행할 필요성이 반영된 것이 GDPR이라는 설명이 있다.(이원복, 앞의 논문, p.199)

GDPR은 신법의 중요한 선례로 아래에서 살펴보는 것과 같이 가명처리(pseudonymisation), 익명처리(anonymization), 암호화(encryption) 등의 개념을 규율하면서 특히 식별가능성에 대한 기준을 적극 제시하고 있다고 평가된다.[93]

GDPR은 개인정보보호에 관한 일반 법규로서 개인정보를 보유하고 있는 기관 등이 개인정보의 수집, 처리 상황을 명확히 인식하고, 개인정보보호책임자(DPO, Data Protection Officer)를 통하여 개인정보보호영향평가(DPIA, Data Protection Impact Assessment) 등을 관리하도록 하며, 식별가능정보(identifiable)에 중점을 두고 정보주체의 다양한 권리를 규정하면서 정보주체의 정보이용에 대한 동의를 요구하면서도, 개인정보의 이전 과정을 명확히 하려는 특징을 보인다고 한다.[94] 특히 GDPR은 기존에 명문으로 규정하지 않았던 가명처리(pseudonymisation)를 상세히 규율함으로써 신법에 큰 영향을 준 것으로 볼 수 있다.

GDPR은 개정 배경을 설명하면서 기존의 정보보호지침이 유럽 연합 내의 통일적인 법적용, 법적 안정성의 보장에 기여하지 못한 점 외에도 온라인에서의 개인정보보호에 대한 광범위한 우려를 해소하지 못한 점을 지적하고, 그러한 문제점으로 인하여 개인정보의 자유로운 흐름(유통)이 방해되어 경제활동의 장애, 경쟁 왜곡 등이 야기되었다고 반성하고 있다.[95]

결국 GDPR은 개인정보의 보호를 개인정보이용에 우선하고자 하는 유럽의 전통을 일응 계승하고 있지만, GDPR 또한 개인정보의 안전한 관리를 통하여 개인정보의 적정한 이용을 활성화하겠다는 문제의식을 본 연구와 공유하고 있다고도 볼 수 있을 것이다. 같은 맥락에서 유럽연합의 관보(Official Journal of the European Union)에 게시된 정식 명칭은 「개인정보의 처리와 자유로운 유통에 관한 자연인의 보호에 관한 규정」으로서 위와 같이 개인정보의 처리, 유통에 관한 고민이 명칭에도 반영되어 있다고 생각된다.[96]

---

93) 양기진, 개인정보의 범위에 관한 연구 - GDPR의 비식별조치와 약학정보원 사건의 검토 -, 선진상사법률연구 통권 제84호, p.62 이하

94) Daniel Solove, "Why I love the GPPR", http://teachprivacy.com/why-i-love-the-gdpr (2022. 7. 17. 확인)

95) GDPR 전문(이하 '전문'이라 한다.) 제9항

GDPR은 총 137개의 조항의 전문(recital)[97]에 이어 총 11장 99개 조문의 본문으로 구성되는데, 개인정보의 처리 원칙, 정보주체의 권리, 개인정보처리자[98], 역외로의 개인정보 이전, 감독기관, 기관간의 관계, 구제절차 등의 내용이 담겨 있다.[99] 이하에서는 본 연구에서 관심을 가지고 있는 비식별조치, 가명처리, 암호화조치 등에 관련된 내용 위주로 정리하고자 한다.

나. 개인정보의 정의와 식별가능성의 문제

규정 제4조 제1항[100]은 식별되거나 식별가능한 자연인을 정보주체(data subject)로 정하면서, 그러한 정보주체와 관련된 일체의 정보를 개인정보로 정의하고 있다. 이 때 특히 「식별가능한 자연인(identifiable natural person)」이라 함은 성명, 식별번호, 위치정보, 온라인 식별자 등과 같은 식별자(identifier) 또는 특정 자연인에 고유한 신체적, 심리적, 유전적, 정신적, 경제적, 문화적,

---

96) “on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC”, 2016. 4. 5.자 Official Journal of the European Union

97) 여러 문헌에서 GDPR의 Recital을 ‘전문’이라고 새기고 있다.[다만, ‘상설(詳說)’이라고 표현하는 문언(박노형. 앞의 책, p.58)과 ‘설명’이라고 처리하는 문헌도 있다. (양기진, 앞의 논문, p.64)] GDPR의 전문은 총 173개의 조문으로 구성되어 있는데 GDPR과 일체로 취급되면서 본문 앞에 첨부되어 있고, 일반적인 전문(前文)에 포함되는 사항, 즉 입법이나 사안의 경위, 배경을 설명하는 부분도 있으나, GDPR에서 사용하는 용어를 정의하는 내용, GDPR의 적용 범위는 물론 정보이용 주체에 대한 의무를 부과하는 규정까지 확인되므로 우리가 이해하는 통상의 전문으로 볼 수는 없다. 또한 위 전문은 GDPR 전체 규정과 함께 유럽연합의 관보에 게재되었고, 이후에도 GDPR과 일체하여 소개되고 있으므로 GDPR 본문과 일체로 규범력이 있다고 봄이 상당하다.

98) GDPR에서는 의무주체를 Controller와 processor로 규정하고 있는데, 직역하면 통제(관리자)와 처리자로 번역할 수 있으나, 규정 전반을 종합할 때 전자는 개인정보를 보관, 보유하고 있는 우리 법상의 개인정보처리자이고 후자는 개인정보처리자로부터 처리를 위탁받은 사람으로 볼 수 있다. 이하에서는 개인정보처리자로 일괄하여 표시하겠다.

99) 이상 L119, 4 May 2016, p.1-88

100) 'personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;

사회적 요소를 통하여 직, 간접적으로 식별될 수 있는 사람으로 정의하고 있어, 식별가능성이라는 요건을 통하여 보호되는 개인정보의 범위를 넓히고 있다고 생각된다.

다. 기타 식별가능성 관련 규정

GDPR은 식별가능성 관련, 식별할 수 있는 정보의 예시를 다양하게 들고 있다. ① 각종 디지털 기기를 통하여 제공되는 인터넷 프로토콜 주소, 쿠키, 주파수 등[전문 제30조, 뒤에서 언급할 부채널 공격(side channel attack)을 염두에 둔 것으로 보인다.]과 ② 안면 영상, 지문과 같은 생체정보[101] 등이 열거되어 있다.

라. 가명처리(pseudonymisation)

GDPR은 비식별조치(de-identification), 익명화(anonymization)라는 표현을 사용하지 않고, 식별된(identified), 식별 가능한(identifiable) 개인정보를 보호하고, 익명처리된(rendered anonymous) 익명정보(anonymous information)를 개인정보의 보호대상에서 제외하고 있는데, 이는 신법이 대체로 계수한 것으로 볼 수 있지만 세부적인 면에서는 차이가 있다. 아래에서는 우선 가명처리 관련 부분을 정리하도록 하겠다.

1) 규정 제4조 제5항, 전문 제26조 및 규정 제32조 제1항 : 가명정보의 개념 및 식별가능성의 판단기준

가) 규정 제4조 제5항

규정 제4조 제5항에서는 가명처리의 개념을 추가 정보의 사용 없이는 정보

101) 규정 제4조 제14호

주체를 식별할 수 없도록 하는 방식으로 개인정보를 처리하는 것이라고 정의하면서, 이러한 경우 추가 정보는 별도로 보관되고, 기술적, 관리적 조치에 따라 정보주체를 식별하는 것에 사용되지 않아야 한다고 정하고 있다. 이는 신법 제2조 제1호의2의 가명처리의 정의 규정, 제28조의4 제1항(추가 정보를 별도로 분리하여 보관, 관리)이 사실상 그대로 계수하고 있는 것이다.

나) 전문 제26조

한편 전문 제26조는 이에 앞서 법의 적용 범위를 「개인정보보호원칙의 범위(the principles of data protection)」로 표현하면서 식별되었거나 식별될 수 있는 자연인과 관련된 모든 정보에 적용한다고 선언하여 일단 개인정보 보호의 범위를 넓히고 있다.

먼저 위 조항은 가명처리(pseudonymisation)를 거친 개인정보라 하여도 추가 정보를 이용하여 자연인을 식별할 수 있다면(be attributed to a natural person by the use of additional information), 자연인을 식별할 수 있는 정보, 즉 개인정보로 보아야 한다고 보아 식별가능성을 가장 중요한 판단기준으로 보고 있고, 신법이 가명정보의 개념을 신설하면서 이를 일단 개인정보에 포함시킨 태도는 위 입법례를 따른 것으로 보인다.

또한 위 조항에서는 식별가능성의 판단 기준을 상술하고 있어 주목하여야 한다. 우선 식별가능성은 정보처리자 또는 제3자가 정보주체를 직·간접적으로 알아보고자 사용할 것으로 합리적으로 예상할 수 있는 모든 수단[102]을 고려하여야 하고, 특히 합리적으로 예상할 수 있는 수단이라 함은 정보를 처리하는 시점 기준으로 사용가능한 기술을 고려하였을 때 정보주체를 식별하는데 드는 시간과 비용과 같은 객관적 요소 전부를 고려하여 판단하여야 한다고 정하고 있다.[103]

---

102) 이와 같은 GDPR의 태도에 대하여 뒤에서 살펴보는 식별주체에 관한 논의에서는 절대설, 상대설 사이의 절충적 입장이라고 평가하는 견해(양기진, 앞의 논문, p.64)가 있고, 사견으로는 GDPR의 취지가 합리적이라고 생각하며 이와 같은 내용으로 입법상의 조치가 필요하다고 본다. 이하 해당 부분에서 상론하겠다.

103) 이상 To determine whether a natural person is identifiable, account should be

우리 신법 및 하위 법규에서는 이와 같은 세부적인 기준을 찾아볼 수 없고, 다만 법 제2조 제1호 나목에서 개인정보에 포함될 수 있는 정보를 판단하는 기준에서 불충분한 형태("이 경우 쉽게 결합할 수 있는지 여부는 다른 정보의 입수 가능성 등 개인을 알아보는 데 소요되는 시간, 비용, 기술 등을 합리적으로 고려")로 확인될 수 있을 뿐이다.

신법의 해석상 적정한 가명처리와 식별가능성의 여부는 동어반복에 해당한다고 볼 수도 있다. 즉, 적정한 가명처리가 이뤄졌다면 그러한 처리로 생성된 가명정보만으로는 정보주체를 식별할 수 없다고 보아야 할 것이기 때문이다.

결국 적정한 가명처리 또는 식별가능성의 유무에 관한 판단기준을 수범자가 충분히 예측할 수 있도록 수립하는 것이 긴요할 것이나, 신법에서는 GDPR과는 달리 가명처리는 물론 식별가능성의 판단기준을 제시하고 있지 않고, 이는 앞서 상세히 지적한 문제점이다.

다) 규정 제32조 제1항

끝으로 위 규정에서는 개인정보의 처리 시, 최신의 기술(the state of the art), 그러한 기술의 실행 비용 및 개인정보 처리의 성격, 범위, 내용 및 목적을 고려하여 적정한 조치를 취하여야 함을 강조하면서 그러한 조치 중 하나로 가명처리 및 암호화조치를 들고 있다. 이러한 규정 또한 적정한 가명처리, 암호화조치를 판단함에 있어 하나의 기준이 될 수 있을 것이다.

2) 전문 제28조 및 전문 제29조

가) 전문 제28조

taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly. To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments.

위 조항에서는 가명처리가 정보주체의 권리를 보호하는 수단임과 동시에 정보이용 주체로 하여금 정보주체의 권리를 보호할 의무를 지원하는 것이라고 강조하는 한편, 가명처리 외에도 다양한 조치를 할 수 있음을 밝히고 있다.[104]

나) 전문 제29조

위 조항에서는 정보이용 주체가 가명처리에 대하여 관심을 가질 수 있도록 하기 위하여, 가명처리를 하는 경우 적정한 가명처리 여부에 대하여 일반적 평가를 허용하되, 동종 업계에서 이 법에서 필요한 조치 및 정보주체에 관한 추가 정보를 분리하는 데에 필요한 기술적·관리적 조치를 취하여야 한다고 규정한다.[105] 이는 신법 제28조의4 제1항이 계수하고 있는 조항으로 보이는데, 특히 GDPR에서는 정보이용 주체가 가명처리를 하는 경우 적정성 평가를 좀 더 용이하게 받을 수 있도록 인센티브를 부여하고 있는 것이 특색이다.

3) 가명처리가 재식별되는 경우를 경계

한편, GDPR은 가명처리의 특성상 역으로 재식별되는 경우를 경계하고 있기도 하다. 예를 들어 전문 제75조에서는 일반적인 가명처리가 재식별되는 경우의 위험성을, 전문 제85조에서는 구체적으로 개인정보가 유출된 경우, 가명처

104) The application of pseudonymisation to personal data can reduce the risks to the data subjects concerned and help controllers and processors to meet their data-protection obligations. The explicit introduction of 'pseudonymisation' in this Regulation is not intended to preclude any other measures of data protection.

105) In order to create incentives to apply pseudonymisation when processing personal data, measures of pseudonymisation should, whilst allowing general analysis, be possible within the same controller when that controller has taken technical and organizational measures necessary to ensure, for the processing concerned, that this Regulation is implemented, and that additional information for attributing the personal data to a specific data subject is kept separately. The controller processing the personal data should indicate the authorized persons within the same controller.

리가 재식별되는 경우의 위험성을 경계하고 있다.

이와 관련하여 신법이 구체적으로 규율하지는 않고 있으나, 특히 전문 제85조의 내용에서 알 수 있듯 불충분한 가명처리된 상태에서 개인정보가 유출된 경우 개인정보의 식별위험성이 중요하므로 이를 적정한 가명처리 판단기준에 반영하는 등의 방식으로 도입할 필요가 있다.

예를 들어 고유식별정보나 비밀번호가 해시함수에 의하여 암호화가 되었지만, 암호화된 해시값이 유출되는 사고가 발생하였을 때, 상대적으로 안전한 해시함수를 사용하지 않거나, 뒤에서 살펴보는 솔트를 추가 정보로 하여 처리하는 등의 추가 조치가 취해지지 않았다면, 공격자는 충분한 시간과 비용으로 이를 복호화할 수 있기 때문에 이러한 위험성을 사전에 방지하는 노력을 취할 수 있도록 이와 같은 가능성, 사례를 가명처리의 적정성 판단 시에 고려할 수 있어야 할 것이다.

4) 규정 제6조 제4항 : 신법 제15조 제3항 등에 따른 이용의 선례

위 규정은 수집 목적과 합리적으로 관련된 범위에서 여러 사정을 고려하여 정보주체의 동의 없이 개인정보의 이용을 허용한 법 제15조 제3항 등에 따른 이용의 선례로 보이는데, 「수집 목적과 추가 처리의 목적 사이의 연관성, 개인정보가 수집된 경위, 개인정보 처리 결과가 초래할 수 있는 결과 및 가명처리, 암호화조치의 존부」 등을 따져 개인정보이용의 허부를 결정하여야 한다고 보고 있으므로, 법 제15조 제3항에 따른 시행령 제14조의2가 위 조항을 사실상 계수하고 있다고 할 수 있다.

마. 암호화조치(encryption)

GDPR은 가명처리와 함께 암호화조치에 관하여도 반복하여 규정하고 있는데, 가명처리와 중첩되지 않는 범위 내에서 정리하여 보면 아래와 같다.

1) 전문 제83조 및 규정 제32조 제1항

GDPR은 전문 제83조에서 개인정보처리자로 하여금 안전성을 확보하고 개인정보의 침해를 방지하기 위하여 정보처리과정에서 내재된 위험을 진단하고, 그러한 위험을 완화시키기 위한 조치를 취하여야 하는데 그 예시로 암호화를 들고 있고.[106] 규정 제32조 제1항에서 개인정보 처리 시의 보안성을 유지하기 위하여 적정한 기술적, 관리적 조치를 취할 것을 규정하고 있는데, 여러 조치 중 하나로 가명처리 및 암호화 처리를 열거하고 있어[107], 개인정보보호의 일반적 수단, 조치의 예시로 암호화를 언급하고 있다.

2) 규정 제6조 제4항

우리 신법 제17조 제4항, 제18조 제3항과 유사한 규정으로, GDPR은 정보주체의 동의 없이 개인정보를 수집한 목적 외로 처리하고자 하는 경우, 수집 목적과의 연관성, 개인정보가 수집된 상황, 개인정보의 성격, 정보주체에 대한 영향에 더하여 가명처리 또는 암호화 여부를 요건으로 하여 허용하고자 하는데[108], 이는 신법 시행령 제14조의2가 참고한 것으로 보인다.

---

106) In order to maintain security and to prevent processing in infringement of this Regulation, the controller or processor should evaluate the risks inherent in the processing and implement measures to mitigate those risks, such as encryption.

107) Taking into account the state of the art, the costs of implementation and the nature, scope, context and purposes of processing as well as the risk of varying likelihood and severity for the rights and freedoms of natural persons, the controller and the processor shall implement appropriate technical and organizational measures to ensure a level of security appropriate to the risk, including inter alia as appropriate:
(a) the pseudonymisation and encryption of personal data;

108) Where the processing for a purpose other than that for which the personal data have been collected is not based on the data subject's consent (중략) the controller shall, in order to ascertain whether processing for another purpose is compatible with the purpose or which the personal data are initially collected, take into account, inter alia (중략 : (a)~(d)) (e) the existence of appropriate safeguards, which may include encryption or pseudonymisation.

3) 규정 제34조 제3항

또한 GDPR은 개인정보 침해가 발생하면 개인정보처리자에게 정보주체에 대한 통지의무를 부과하고 있으나, 위와 같은 적정한 기술적, 관리적 조치를 취한 경우에는 면제하면서, 특히 권한 없는 제3자가 개인정보를 이해할 수 없도록 하는 암호화와 같은 조치를 취한 경우를 예로 들고 있다.[109]

바. 익명정보(anonymous information)

GDPR은 위 전문 제26조에서 익명정보에 대하여는 개인정보보호의 원칙을 적용하지 않으므로 GDPR이 적용되지 않는다고 규정하면서, 익명정보를 식별되거나 식별가능한 자연인과 관련되지 않는 정보 또는 정보주체가 더 이상 식별되지 않을 정도로 익명처리된 경우의 정보로 정의하고 있다. 다만 GDPR은 구체적인 익명처리 방법을 언급하고 있지는 아니하므로 어떠한 처리든 간에 정보주체가 식별되지 않는다면 익명정보로 볼 수 있을 것인데, 이는 신법 제58조의2가 적절히 규율하고 있는 것과 동일하다.

사. 검토

이상 종합하면, GDPR은 신법과는 달리 적정한 가명처리, 암호화조치의 판단에 있어 세부적인 기술적 사항을 들지 않고도, 규범적 판단을 내릴 수 있는 일응의 기준을 제시하고 있다고 할 수 있다.

GDPR은 ① 식별주체를 개인정보처리자 외에 제3자로 명확히 하고, ② 그들

---

109) 3. The communication to the data subject referred to in paragraph 1 shall not be required if any of the following conditions are met:
(a) the controller has implemented appropriate technical and organizational protection measures, and that those measures were applied to the personal data affected by the personal data breach, in particular those that render the personal data unintelligible to any person who is not authorized to access it, such as encryption;

이 사용할 것으로 합리적으로 예상할 수 있는 모든 수단의 고려하도록 하면서, 그러한 수단은 ③ 정보를 처리할 때 사용가능한 기술(즉, 최신의 기술), 정보주체를 식별하는데 드는 시간과 비용 등을 감안하여 범위를 제한 할 수 있고, ④ 이에 더하여 개인정보 처리의 성격, 범위, 내용 및 목적을 종합하는 방법으로 가명처리 등의 적정성을 판단하여야 한다고 규정하고 있다.

또한 GDPR은 정보이용 주체의 정보 처리와 정보주체의 보호의 조화를 강조하고 있고, 기술적으로는 개인정보가 유출되었을 때의 위험성까지 고려하여 가명처리 등 안전성 확보 조치를 취하여야 함을 환기시켜주고 있다고 본다. 이에 대하여는 GDPR이 개인정보 보호를 위한 비식별조치의 일반적인 방안으로 가명처리를 받아들이고 있기 때문으로, 가명처리를 개인정보의 이용에 관한 하나의 특례로 보고 있는 우리 신법의 태도와 구별된다고 하겠다.[110] 덧붙여 GDPR은 가명처리와 함께 암호화조치를 반복하여 열거하는데, 여기에서 암호화조치는 일반적인 안전성 확보를 위한 기술적, 관리적 조치의 대표적인 예시로 등장하고 있다.

요컨대, GDPR이 적정한 가명처리의 판단기준에 대하여 상세히 규정하고, 개인정보의 보호 수단으로서 가명처리를 적극 내세우고 있음은, 개정 과정에서 GDPR을 일부 계수한 것이라고 평가받고 있는 개정 개인정보보호법에서 찾아볼 수 없는 것으로 이와 같은 사정은 본 연구에서 제기된 입법 조치의 미비를 해결하는 하나의 열쇠가 될 수 있을 것이다.

2. 미국 : 의료보건 데이터의 보호와 활용

미국의 경우, 정부가 국가안보의 측면에서 암호화 알고리즘의 개발을 주도, 관리해 온 역사적 경험이 있다. 그 과정에서 정부는 컴퓨터 연산, 수리암호관련 민간의 기술이 발달하게 되자 오히려 민간의 암호화조치를 억제하려고 노력한 적이 있었고, 이에 대하여 기업 등 시민들은 정부의 개입이 수정헌법 제1조가 보장하는 표현의 자유, 사생활의 자유를 침해한다는 취지로 강력히 반

110) 이원복, 앞의 논문, p.200

발한 사례가 있어 가명처리나 암호화조치에 대한 정부 정책의 방향과 내용이 EU나 우리와는 사뭇 다르다. 한편으로 미 연방대법원은 정보이용 주체의 개인정보이용을 제한하는 것은 위와 같은 표현의 자유를 부당하게 제한하는 것이라고 보기도 하였다.

이와 같이 미국은 정부가 정보이용 주체에 대하여 어떠한 의무를 부과하는 것에는 상당한 거부감이 있는 것으로 보이는데 이는 고유한 역사적 전통에 따른 것이 아닌가 한다.

그러나 한편으로는 의료보건 데이터와 같이 우리 법체계에서 민감정보로 규율하는 정보에 대하여는 HIPAA와 같은 적극적인 입법을 통하여 정보주체의 권리 보호에도 관심을 가지고 있다. 이하에서 상세히 살펴보도록 한다.

가. 개인정보이용과 표현의 자유

1) 암호화조치에 대한 정부와 민간의 태도

아래에서 편을 바꾸어 살펴볼 다양한 암호화 알고리즘은 사실상 미국이 정부 주도로 공모, 채택하는 등으로 개발된 DES, AES, SHA와 같은 수리암호화 주축을 이루고 있다. 학문적, 기술적 우위를 바탕으로 정부가 암호화 기술을 선도하여 온 셈이다. 다만, 그 과정에서 1980년대 이전까지는 정부가 암호화 알고리즘을 기밀의 암호화, 복호화, 정보수집 등 국가안보의 목적에서 주로 사용한 것으로 볼 수 있다.[111]

그 과정에서 정부는 퍼스널 컴퓨터, 인터넷이 발달하고 원격지 간의 통신, 국제 금융거래 등에서 암호화조치를 이용한 처리가 활성화되자, 정부 차원에서 국가 안보, 범죄 예방, 수사 등의 목적을 내세워 암호화 알고리즘 관련 소프트웨어의 수출을 통제하거나, 암호화키를 정부기관에 맡기는 소위 키 에스크로(key escrow) 정책 내지 일정한 조건을 충족하면 임의로 복호화할 수 있는 소위 백도어(back-door)기능을 도입하려고 하는 시도를 하게 되었다.[112]

---

111) Jan H. Samoriski 등 3명, Encryption and the first amendment(Mar. 23rd, 2009), 2 COMM. L. & POL'y 417 at 422, 423

그러나 이와 같은 정부의 시도는 입법화 되지도 못하고 의회의 단계에서 실패하였는데, 가장 큰 이유는 사인이 통신 과정에서 암호화조치를 취하는 것은 수정헌법 제1조가 보장하고 연방대법원이 강력히 보호하고 있는 표현의 자유(freedom of speech)의 범위 내에 포섭되는 것으로서 정부가 아주 예외적이고 절박한 이익(compelling interest)을 입증[113]하여야 했기 때문이다.[114]

결국 미국의 경우 민간은 스스로의 사생활 보호, 기업활동의 자유를 위하여 암호화조치를 취하면서 표현의 자유를 내세웠고, 정부는 민간의 암호화조치가 정부의 통제로 벗어나는 것을 우려하였다고 볼 수 있어, 가명처리나 암호화조치를 개인정보처리자에 대한 의무로 취급하는 EU나 우리의 현실과는 구분할 필요는 있다.

2) Sorrel v. IMS Health Inc. 사건

IMS Health는 아래에서 살펴볼 '약학정보원 사건'에서 처방정보를 구입한 회사의 미국 소재 모회사이다. 위 회사는 1954년 설립되어 미국 코네티컷 주 댄버리 등에 본거지를 둔 회사로 전 세계 100여개 국가에서 수집하는 의학 관련 국제통계에 기반을 둔 의료정보제공, 기술개발에 주된 사업을 영위하였는데, 2017. 11. 6. 아이큐비아(IQVIA)로 사명을 변경한 회사이다.

그런데 위 회사는 약국으로부터 처방자(의사)가 식별되는 정보를 수집하고 이를 정보처리업자(data miners)에게 판매하는 업을 하였고, 정보처리업자들은 위 처방정보를 토대로 의사의 처방 습관에 대한 보고서를 제약회사에 판매하

112) 2 COMM. L. & POL'y 417 at 435, 438

113) 인터넷 통신이 대중화 되던 초기 미 연방의회가 통신품위법(the Communications Decency Act of 1996, CDA)을 제정하여 성적인 표현 등을 미성년자에게 노출하는 등의 정보를 제한하고자 하였으나, 연방대법원은 이와 같은 법률은 표현의 자유를 침해하여 위헌이라고 보았는데 마찬가지 판단 기준을 적용하였다. [Reno v. ACLU, 512 U.S. 844(1997)]

114) 미 연방정부는 이후에도 이와 같은 시도를 계속한 것으로 보이나 결국은 실패하였다. (Jeri Clausing, U.S. Losing Battle on Control of Data Encryption, Study says, Feb. 9, 1998, The New York Times, Joel Brinkley, U.S. Eases Encryption Software Export Bans, Sep. 17, 1998, The New York Times) 그러나 주지하다시피 2013년 미국 국가안보국(National Security Agency, NSA)의 각종 백도어 프로그램 개발, 이용을 스노든(Edward Snowden)이 폭로하면서 연방정부의 시도는 그 이후에도 암암리에 계속된 것으로 보인다.

며, 제약회사는 이틀 토대로 제약회사가 의사들에게 의약품을 판매할 때 세부 설명(detailing)을 제공한다.

이에 대하여 버몬트 주정부는 2007년 제정한 처방비밀보호법(Prescription Confidentiality Law)으로 의사의 동의 없이 의사의 처방내역을 판매, 공개하는 등 상업적 용도로 사용할 수 없도록 하였고, 이에 대하여 위 회사는 위 법의 적용을 막고자 법에 대한 가처분(injunctive relief)을 구하였는데, 이에 대하여 버몬트 연방지방법원은 가처분 신청을 기각하였지만, 연방 제2항소법원은 이를 인용하여 버몬트 주 법무장관 Sorrell이 연방대법원에 상고하게 되었다.[115]

이에 대하여 연방대법원은 정보를 생성, 보급하는 것도 표현의 자유의 범위 내에 속하고, 위 법이 정보수집자(data miners)에 대하여 그러한 의사표현의 내용, 여부 등에 대하여 구체적인 제한을 가하고 있는 반면, 의료인의 비밀유지의무, 의사-환자와의 관계 등에 관한 주 정부의 이익만으로는 그러한 제한을 정당화 할 수 없다고 보았다. 세부적으로는 의약품 판촉을 위한 의사표현도 표현의 자유로 보호되는 것이어서 엄격한 심사척도(strict scrutiny)가 적용되는데 주 정부가 의사의 처방내용이 모니터링된다는 반발, 환자들의 우려 등 입법의 절박한 이유를 제대로 입증하지 못하였다는 것이다.[116]

신법이나 GDPR의 현행법에 따르면 위와 같은 처방 정보로 환자는 물론 의사를 식별할 수 있는 이상, 그 처리에 엄격한 제한이 따를 것으로 보이는데 미 연방대법원은 수정헌법 제1조에서 보호되는 표현의 자유를 넓고, 강력하게 보호하고 있는 전통에 따라 정보를 생성, 보급하는 것도 표현의 자유라고 보아 이를 규제하는 법률에 대해 엄격한 심사척도를 적용하고 있는 셈이다.

즉, 미국의 경우 입법자가 우리의 개인정보보호법이나 GDPR과 같은 입법으로 개인정보처리자의 처리를 제한하려면 정부가 입법 목적이 중대하고 절박한

115) 이와 유사한 입법이 메인 주, 뉴햄프셔 주 등에도 있었고, 전국적 영업을 하던 IMS Health는 각 주 정부를 상대로 유사한 소송을 제기하였으나, 각 연방항소법원들은 회사의 정보 수집을 제한하는 것이 행동(conduct)에 대한 제한으로 표현의 자유에 포섭되지 않는다고 보아 회사의 신청을 기각한 상태였다.(IMS Health Inc. v. Mills, 616 F.3d 7 및 IMS Health Inc. v. Ayotte, 550 F.3d 42)

116) Sorrel v. IMS Health Inc. 564 U.S. 552, at 556, 557, 558

공익과 관련된 것임을 확실히 입증하여야 할 것이고, 미 연방대법원의 그간의 판례에 따른다면 일단 표현의 자유의 범위 내에 포섭되는 순간 이를 제한하는 입법이 합헌성을 확인받는 것은 극히 예외적일 것이다.

이러한 법현실 속에서도 아래에서 살펴보는 미 연방 법률인 의료정보보호법은 의료보험에 포함된 정보의 민감성에 주목하여 환자의 개인정보를 보호하기 위하여 비식별조치를 강력히 요구하고 있어 살펴볼 필요가 있다.

나. 연방 의료정보보호법(HIPAA)

또한, 미국 연방 법률인 HIPAA의 개인정보규칙(Privacy Rule)에서는 '비식별조치'로 직역될 수 있는 'de-identification'의 개념을 사용하고 있어 검토할 필요가 있다.

우선 위 법에서는 '개인이 식별될 수 있는 의료정보(individually identifiable health information)'를 「의료보험사업자 등이 수집, 생성하거나, 개인에 관한 (현재, 과거, 미래의) 신체, 정신의 건강 등 관련 정보, 의료보험의 내용, 의료보험료의 지급정보로서, ① 개인을 식별할 수 있거나, ② 그 정보가 개인을 식별하는데 사용될 수 있다고 믿을 만한 합리적 근거가 있는 경우[117]」로 정의하였고, 이러한 정보와 관련된 건강 식별자를 사용하거나 사용되게 한 경우, 이러한 정보를 취득, 공개한 경우를 강력하게 처벌하고 있다.[118]

---

117) The term "individually identifiable health information" means any information, including demographic information collected from an individual, that--
(A) is created or received by a health care provider, health plan, employer, or health care clearinghouse; and
(B) relates to the past, present, or future physical or mental health or condition of an individual, the provision of health care to an individual, or the past, present, or future payment for the provision of health care to an individual, and--
(i) identifies the individual; or
(ii) with respect to which there is a reasonable basis to believe that the information can be used to identify the individual.
[이상 42 U.S.C.A. § 1320d (West)]

118) 기본적 구성요건에 해당할 경우 5만 달러 이하의 벌금, 1년 이하의 징역의 형사처벌이

계속하여 법률의 시행규칙 격인 HIPAA 개인정보보호규칙(Privacy Rule)에서는 법률의 내용을 상세히 규율하고 있다.

우선 개인적으로 식별 가능한 의료정보를 정보주체로부터 취득한 신상정보가 포함된 의료정보의 일체라고 정의하면서, ① 그러한 정보는 의료보험사, 직원 등에 의하여 생성되는 것으로, 현재, 과거, 미래의 신체, 정신 건강 또는 개인의 상태에 관련된 것 또는 건강보험 급여 지급 내역 등에 관한 내용에 의하여 정보주체가 식별 가능한 것이거나, ② 합리적인 이유에 따라 그러한 정보가 개인을 식별하는데 사용할 수 있을 것으로 여겨지는 것[119]이라고 하여 법의 규정을 좀 더 상술하되 식별가능성을 중요한 기준으로 삼고 있다.[120]

그 다음으로 위 규정은 「보호되는 건강정보의 비식별조치(De-identification of protected health information)」라는 표제 하에, 건강정보 중 개인을 식별하지 못하거나 개인을 식별하는데 사용될 수 있다고 믿을만한 합리적인 근거가 없는 경우에는 '개인이 식별될 수 있는 건강정보'에서 제외한다고 규정[121]하면

규정되어 있으나, 사위행위로 인한 범행의 경우 10만 달러 이하의 벌금(병과 가능), 5년 이하의 징역, 이를 판매, 배포할 목적을 가지고 범행한 경우 또는 상업적 이익이나 악의적 행위를 위하여 사용한 경우에는 25만 달러 이하의 벌금, 10년 이하의 징역의 각 가중처벌을 규정되어 있다.[각 벌금형 선택 또는 병과 가능, 이상 42 U.S.C.A. § 1320d-5 (West)]

119) 이는 결국 정보처리자가 당해 정보에다가 합리적으로 접근할 수 있는 '다른 정보'와 결합하여 정보주체를 식별할 수 있는지 여부라고 설명하는 견해도 있다.(이인호, 앞의 논문, p.80)

120) Individual means the person who is the subject of protected health information.
Individually identifiable health information is information that is a subset of health information, including demographic
information collected from an individual, and:
(1) Is created or received by a health care provider, health plan, employer, or health care clearinghouse; and
(2) Relates to the past, present, or future physical or mental health or condition of an individual; the provision of health care to
an individual; or the past, present, or future payment for the provision of health care to an individual; and
(i) That identifies the individual; or
(ii) With respect to which there is a reasonable basis to believe the information can be used to identify the individual.
[이상 45 C.F.R. § 164.103]

121) a) Standard: De-identification of protected health information. Health information

서 이를 판단하는 기준을 제시하고 있다.[122]

첫째, 적정한 지식과 경험, 통용되는 통계학적, 과학적 기법을 사용할 수 있는 사람('전문가')이, 정보 그 자체로, 또는 합리적으로 사용가능한 정보를 결합하여도 정보주체를 식별할 가능성이 매우 낮다는 의견을 개진하면 비식별조치를 인정할 수 있고, 둘째, 정보주체 또는 그의 친척, 고용주, 동거인에 관한 식별자(identifier)를 삭제한 경우에도 마찬가지라고 한다. HIPAA는 위와 같은 식별자에 해당하는 정보로 성명, 전화번호, 주 단위 미만의 주소, 도시, 카운티, 우편번호(첫 번째 3자리 제외)로서 인구조사국에서 확인가능한 정보, 생년월일, 입대-제대일, 사망일, 전화번호, 이메일 주소, 사회보장번호, 운전면허 번호 등이 열거하고 있기도 하다.[123]

다. 검토

---

that does not identify an individual and with respect to which there is no reasonable basis to believe that the information can be used to identify an individual is not individually identifiable health information. [이상 45 C.F.R. § 164.514]

122) 같은 취지에서 비식별조치 등 규정에 따라 처리된 의료정보는 더 이상 식별 가능한 의료정보로 보지 않는다고도 한다. [이상 45 C.F.R. § 164.502]

123) (b) Implementation specifications: Requirements for de-identification of protected health information. A covered entity may determine that health information is not individually identifiable health information only if:
(1) A person with appropriate knowledge of and experience with generally accepted statistical and scientific principles and methods for rendering information not individually identifiable:
(i) Applying such principles and methods, determines that the risk is very small that the information could be used, alone or in combination with other reasonably available information, by an anticipated recipient to identify an individual who is a subject of the information; and
(ii) Documents the methods and results of the analysis that justify such determination; or
(2) (i) The following identifiers of the individual or of relatives, employers, or household members of the individual, are removed:
(A) Names; (중략) (D) Telephone numbers; (중략) (F) Electronic mail addresses; (G) Social security numbers; (중략) (K) Certificate/license numbers;(후략) [이상 45 C.F.R. § 164.514]

이상 간략히 살펴본 것과 같이 개인의 표현의 자유를 중시하는 미국에서도 민감정보 중 하나인 의료정보에 대하여는 연방법으로 상세히 규율하여 이를 보호하고 있다. 특히 HIPAA에서는 적정한 비식별조치가 인정된 경우에는 보호대상인 의료정보에서 제외하면서 신법의 익명정보와 유사한 취급을 하고 있는데 신법에 비하여 비식별조치, 식별자의 개념을 명확하게 정하고 있다고 판단된다.

살펴본 것과 같이 HIPAA는 비식별조치가 적정히 되었는지 여부에 대하여는 2가지 방식을 택하고 있는데, 첫 번째 방안은 적정한 지식과 경험, 통용되는 통계학적, 과학적 기법을 사용할 수 있는 전문가가 판단하는 것으로 하면서 그 세부적인 기준을 상세히 정하고 있다. 이는 비록 HIPAA가 명시적으로 가명처리나 암호화조치를 요구하고 있지는 않으나, 아래 두 번째 방안이 식별자를 삭제하는 방식을 취하고 있으므로 식별자의 삭제 이외의 비식별조치를 받아들이는 것으로 볼 수 있고, 그 판단기준을 상세히 규율함으로써 예측가능성을 높이고 있다고 평가할 수 있다.

두 번째 방안은 정보주체, 친척, 사용인, 동거인 등의 열거된 개인정보가 삭제된 경우 비식별조치가 되었다고 보는 방식으로 신법 상 가명처리의 삭제, 마스킹에 해당한다고 볼 수 있다.

요컨대 HIPAA도 GDPR과 마찬가지로 법령의 단계에서 '합리적인 이유에 따라 그러한 정보가 개인을 식별하는데 사용할 수 있을 것으로 여겨지는 것'이라고 상술하면서 그 판단기준을 상세히 정리하고 있는 것은 GDPR과 같은 맥락에서 시사하는 바가 크다고 하겠다.

Ⅳ. 적정한 비식별조치로서의 가명처리

개정 개인정보보호법에서 도입된 가명처리는 구법 시행 당시부터 논의되었던 비식별조치의 하나라고 볼 수 있다. 이하에서는 본 연구에서 현재까지 진행된 논의를 종합하여 비식별조치라는 개념의 유래, 가명처리, 암호화조치의 개념을 정리한 후, 기존 판례의 해석론에서 나타나는 문제점을 확인한 다음 연구 결과를 종합하여 입법론, 법률의 해석 및 적용과 기술적 측면에서 암호화조치의 적극적 활용에 관하여 생각해 보겠다.

1. 비식별조치

구법이 시행되던 시기부터 강학상 비식별조치(de-identification)에 대한 논의가 있어 왔고, 이는 GDPR, HIPAA와 같은 외국 법제에서 개인을 식별할 수 없도록 하는 조치로서 규정되어 온 것임은 이미 정리해 보았다. 한편 구법에서도 '익명처리(구법 제3조 제7항)'의 개념이 규정되어 현재까지 이르고 있다. 이하에서는 비식별조치가 어떠한 의미인지, 신법상의 가명처리, 암호화조치는 어떻게 새겨야 할 것인지 검토한다.

가. 구법상 개인정보 비식별조치 가이드라인

구법에서는 가명정보, 가명처리의 개념이 없었으므로, 개인정보는 '개인을 알아볼 수 있는 정보' 또는 '해당 정보만으로는 특정 개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 알아볼 수 있는 것'으로 정의(구법 제2조 제1호)되고 있었다.

또한 구법에서는 정보주체의 동의와 수집 목적 내의 이용을 원칙으로 삼으면서도 구법 제18조 제2항 각호의 사유가 인정되면 예외적으로 개인정보의 목적 외 이용, 제공을 허용하고 있었는데, 그 중 제4호가 「통계작성 및 학술연구 등의 목적을 위하여 필요한 경우로서 '특정 개인을 알아볼 수 없는 형태'로 개

인정보를 제공하는 경우」를 들고 있어, 이 때 '개인을 알아볼 수 없는 형태'가 어떠한 요건에 따라 인정되는지에 대하여 논의가 있어왔다.

결국 구법의 해석으로도 정보주체를 식별할 수 없다면, 일정한 요건에 따라 개인정보의 이용이 제한적으로 가능하다거나, 개인정보로서 보호되지 않으므로 정보주체의 동의 없이 처리할 수 있다는 입론이 가능하였는데, 문제는 법률에서 정보주체를 식별할 수 없도록 하는 조치, 즉, 비식별조치에 관하여 개념, 요건을 다루지 않았다는 것이다.

다만 구법 당시 행정자치부, 방송통신위원회 등 7개 유관기관이 합동으로 비식별조치에 대한 일응의 기준으로 제시하기 위하여 '개인정보 비식별조치 가이드라인'을 제작, 공표하기에 이르렀고, 동 가이드라인은 비식별조치가 적정한 수준에 이르면 위 구법 각 조항의 제한, 즉 보호받은 개인정보에 해당하지 않을 수 있다고 설명하였다.

그러나 위 가이드라인은 모법인 구법에서 '비식별조치'가 규정되지 않은 상황에서 입법 조치 없이 무리하게 개인정보의 예외적 이용을 유도하려는 시도로 비판의 대상이 된 점, 위에서 이미 살펴본 것과 같이 가이드라인이라는 형식의 행정행위가 지닌 태생적 한계로 논란이 야기된 점[124], 비식별조치에 대한 이해가 부족한 모습을 일부 드러낸 점[125], 실제 아래에서 살펴보는 것과 같

---

124) 가이드라인이 상위법의 위임 근거가 없으므로 기속력이 없었다는 지적도 같은 맥락에서 이해할 수 있다. (양기진, 앞의 논문, p.73) 다만, 통상 기속력이라는 개념은 국가권력이 법률, 판례 등의 취지에 따르게 하는 효과이므로 본 연구에서는 수범자, 사법기관에 대한 효력인 규범력이라고 표현하겠다.

125) 예를 들어 개인정보 비식별조치 가이드라인에서는 "비식별조치는 EU의 익명화와 사실상 같은 개념입니다."라는 설명(국무조정실 등 4개 부처, 개인정보 비식별조치 가이드라인(2016), p.7)이 있는데, 이는 '가명처리'의 개념이 도입되어 익명정보와 가명정보가 구별되어야 하는 신법에 대한 해석에서는 원용할 수 없는 것이다. 우선 구법에서도 신법과 동일하게 개인정보를「해당 정보만으로는 특정 개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 알아볼 수 있는 것」이라고 정의하였는데, 비식별조치가 암호화조치로 이뤄진 경우에는 복호화 등 특별한 조치가 있다면 개인정보의 재식별이 충분히 가능하다. 따라서 비식별처리를 익명화와 동일한 것으로 보겠다는 설명은 중대한 오류이다. 만일 위 가이드라인이 비식별처리를 통하여 완전한 익명화를 요구한 것이라면 상당수의 경우에 개인정보의 이용이 불가능하게 될 것이다. 이에 따라 신법은 위와 같은 문제점을 입법으로 해결하고자 '익명정보'와 '가명정보'를 구별하여 상세히 규정한 것으로 보인다.(이를 지적한 견해로는 박노형·정명현, 빅데이터 분석기술 활성화를 위한 개인정보보호법의 개선 방안 - GDPR과의 비교 분석을 중심으로, 고려법학 제85권, p.30)

이 비식별조치가 되었음을 내세우면서 정보주체의 동의 없는 개인정보 처리를 시도한 사례에서 실질적으로는 부적정한 비식별조치가 있었던 점 등에서 오히려 데이터 3법의 개정 계기가 되었다고 보는 것이 온당하다.

나. GDPR의 비식별조치

비식별조치는 우리 법률상 정해진 개념이 아니고, 이를 정면에서 논한 판례를 찾기도 어려운 것으로 국내에서는 구법 시행 당시의 위 가이드라인에서 위와 같이 처음으로 공적인 취급을 받게 되었다.

한편, 이미 수차례 언급한 것과 같이 우리 정보보호법제에 중요한 선례에 해당한다고 평가되는 GDPR에서는 가명처리(pseudonymisation), 익명처리(anonymization), 암호화(encryption) 등의 개념을 규율하면서 특히 식별가능성에 대한 기준을 적극 제시하고 있다.[126)]

특히 앞서 살펴본 GDPR은 제4조 제5항에서 가명처리를 「추가 정보를 사용하지 않는 한 더 이상 특정 정보주체를 알아볼 수 없도록 하는 개인정보의 처리 방식으로서 그러한 추가 정보는 분리 보관되는 한편, 기술적, 조직적 조치를 통하여 자연인을 식별하거나 식별할 수 있는 것에 사용되지 않도록 관리되어야 한다.」고 규정하였고, 전문 제26항에서는 「익명정보(anonymous information)에는 개인정보보호의 원칙이 적용되지 않는다. 익명정보라 함은 자연인을 식별하거나 식별할 수 있는 정보와 관련되지 않는 것이거나, 정보주체가 식별되지 않거나 더 이상 식별되지 않는 방식으로 익명처리된 개인정보를 의미한다.」고 정의하고 있어, 식별가능성을 준거로 하여 개인정보의 보호는 물론 개인정보의 적정한 이용을 도모하는 개인정보보호법제의 모델을 제시하였다고 평가할 수 있다.

다. 검토

126) 양기진, 앞의 논문, p.62 이하

이상 살펴본 구법 하의 해석론, GDPR의 규정 취지를 종합하면, 비식별조치는 식별가능성을 없애거나 식별가능성 내지 위험성을 현저히 낮게 하는 조치로서, 완전히 정보주체를 식별할 수 없게 하는 조치뿐만 아니라 다른 정보와 결합하는 등의 여러 가지 사정으로 식별가능성이 남아 있더라도 그 식별가능성을 일정한 요건에 따라 차단하는 조치로 정의할 수 있다고 생각한다.

결국 비식별조치를 신법상의 개념으로 정리하면 위 전자는 익명처리에 해당한다고 볼 수 있고, 후자는 정확히 일치하지 않지만 가명처리와 일맥상통한다고 할 것인데[127], 실제 GDPR의 규정을 입법 과정에서 참고한 것으로 알려진 신법에서의 가명처리 규정은 GDPR에서의 가명처리(pseudonymisation)와 유사하다고 평가되고, GDPR의 경우 법률 및 하위 법규의 단계에서 비식별조치의 판단 기준을 비교적 상세히 정한 것으로 보인다.

그럼에도 우리의 신법에 따른 각종 법령 등은 적정한 비식별조치에 대한 판단 기준을 충분히 제시하고 있다고 보기 어려운 점은 이미 논증한 것과 같으므로 아래에서는 가명처리 및 암호화조치에 관하여 좀 더 깊게 살펴보아 앞으로의 개선책을 강구하도록 하겠다.

2. 가명처리와 암호화조치

가. 가명처리로서의 암호화조치

신법에서는 명시적으로 가명처리의 개념을 도입하여 개인정보의 일부를 삭제하거나, 일부 또는 전부를 대체하는 등의 방법으로 원래의 상태로 복원하기 위한 추가 정보의 사용·결합 없이는 특정 개인을 알아볼 수 없도록 처리하는 것」을 가명처리로 정의하고 있음은 이미 정리하였다.

한편 구법에서부터 암호화를 '안전성 확보에 필요한 조치'의 예시로 들고 있고(법 제24조 제3항), 신법에서는 아래에서 언급하는 것과 같이 일정한 요건에

127) 실무에서는 신법 시행 이후에도 여전히 익명처리와 가명처리 양자를 포괄하여 비식별처리의 개념을 원용하고 있다고 소개하는 견해가 있다. [고학수 등 7인, 앞의 책, p.5]

따라 정보주체의 동의 없는 개인정보의 처리, 제공을 허용하게 되었는데 그 때에도 암호화 등 안전성 확보에 필요한 조치를 요건으로 삼고 있다.(법 제15조 제3항, 제17조 제4항)

그러나 신법에서 위와 같이 가명처리의 개념을 정하였음에도 더 나아가 적정한 가명처리를 판단할 수 있는 요건을 다루는 것에 이르지 못하였고, 그와 같은 한계를 시행령에서도 그대로 노정되고 있어, 결국 가명처리에 대한 상세한 설명은 개보위의 가이드라인에서 다루고 있음을 살펴보았다.

신법, 하위 법령은 가명처리와 암호화조치를 병치하여 규정하거나 양자의 관계를 명시적으로 정하고 있지는 않으나, 위 개보위의 가이드라인이나 일반적인 견해는 가명처리의 한 가지 형태인 「개인정보 등의 삭제 또는 대체」의 기술적 방법의 하나로 암호화조치를 들고 있다.[128)]

아래에서 상세히 살펴보는 것과 같이 단순한 암호(cryptography)가 아니라 암호화조치(encryption)인 이상, 특정한 정보가 없는 사람은 평문을 알아볼 수 없게 하되, 안전한 조치에 따라 특정한 정보를 보호하여 안전성을 높이는 암호화조치야말로 가명처리의 전형이라 할 수 있으므로 암호화조치는 가명처리의 한 방편이라고 봄이 상당하다.

한편, 현대 산업사회에서 개인정보처리자가 개인정보를 처리할 때에는 처리되는 개인정보의 양, 처리 속도, 처리 결과의 이용의 측면에서 디지털데이터의 형식으로 컴퓨터 등 전산기기를 이용하여 처리하는 것이 일반적이고, 이와 같은 개인정보의 처리 과정은 디지털데이터의 전송, 연산, 저장, 출력 등 다양한 형태로 빈번하게 이뤄짐은 다시 강조할 필요가 있다.

위와 같은 개인정보 처리의 전 과정에서, 우선 개인정보 보호의 측면에서는 '해커'와 같은 외부자의 공격과 개인정보의 무단 유출과 같은 개인정보처리자 또는 개인정보처리자의 임직원 등 내부자의 비행으로 인한 위협이 예상되고, 신법이 단순히 개인정보의 안전한 보관을 넘어 효과적인 이용까지 기도하고 있으므로, 결국 개인정보를 안전히 보호함과 동시에 이를 활용하는 것이 중요한 과제[129)]라고 할 것이다. 결국 본 연구의 궁극적 목적은 적정한 가명처리,

128) 예를 들어 고학수 등 7인, 앞의 책, p.57 이하
129) 신법 개정 이전 구법 당시의 비식별조치 가이드라인을 비판하는 견해는 "최고 수준으로

암호화조치가 위와 같은 과제 해결의 전제라는 것을 논증하고, 이를 위하여 암호화조치를 포함한 기술적 고려를 가미하여 각종 규범의 보완, 법의 해석, 적용에 대한 새로운 기준을 제언하는 것이라고 하겠다.

이하에서는 앞서 개보위의 가이드라인에서 구체적으로 정리하지 않은 암호화조치의 개념, 특성에 대하여 개인정보의 보호와 이용, 두 가지의 측면에서 살펴보겠다.

나. 암호화조치의 개념

암호화(Encryption)의 사전적 의미는 메시지를 암호문(cipher)로 전환하는 것[130]으로, 암호화를 이용한 비밀 정보의 전달 과정은 정보를 보내려는 송신자가 평문(plain-text) 상태의 비밀 정보를 암호화 알고리즘에 따라 암호화키(encryption key)를 사용하여 암호문(cipher)로 변환한 후, 이를 수신자에게 전달하면, 수신자는 복호화 알고리즘에 따라 복호화키(decryption key)를 사용하여 암호문을 복호화하여 평문을 알아내는 과정을 거치게 된다.[131]

이와 같은 암호화 과정을 이용한 정보의 전달에 있어, 외부자, 내부자와 같은 도청자(eavesdropper)가 암호문을 알더라도 평문을 알 수 없는 기밀성(secrecy), 수신자가 받은 암호문을 복호화한 평문이 송신자의 평문과 동일하다는 무결성(integrity), 위와 같은 과정을 적당한 시간과 비용에 따라 용이하게 구현할 수 있어야 한다는 효율성(efficiency)은 적정한 암호화조치의 판단에 기본적 요소이다.[132]

(데이터를) 세탁하면 안전하겠으나, 데이터로서의 가치가 떨어진다. 반대로 그(비식별화) 수준을 낮추면 낮출수록 데이터의 가치는 높아지겠으나, 재식별의 위험이 증가한다."는 설명을 하였다.(오길영, 앞의 논문, p.341) 그러나 신법이 시행된 이상 신법의 적정한 해석, 적용을 통하여 그 접점을 규범적으로 모색할 단계가 되었다.

130) https://www.merriam-webster.com/dictionary/encryption(2021. 11. 1. 확인) 한편 암호를 의미하는 cryptography는 그리스어에서 유래된 단어로 'hidden writing;을 의미한다고 한다.[Laurence D. Smith, Cryptography, The science of Secret Writing(1942) Jan H. Samoriski 등 3명, 앞의 논문, 2 COMM. L. & POL'y 417, at 419에서 재인용]

131) 김명환, 앞의 책, p.3

132) 김명환, 앞의 책, p.4 그에 더하여 위와 같은 암호의 특성을 응용한 기술로는 인증과 서명(authentication and signature), 비밀 분산(secret sharing), 협동 계산(multiparty

이와 같은 암호화조치에 대하여는 해커, 내부자 등 공격자가 가지고 있는 정보량에 따라, 공격자가 몇 가지 암호문을 가지고 있는 경우, 공격자가 특정 평문에 대응하는 특정 암호문의 쌍을 몇 가지 알고 있는 경우, 공격자가 자신이 임의로 선택한 몇 개의 평문을 암호화할 수 있거나, 반대로 몇 개의 암호문을 평문으로 복호화할 수 있는 경우를 가정하고, 그 안전성을 따져볼 수도 있을 것이다.[133]

결국 개인정보의 보호, 처리의 실무에서는 불충분한 암호화 및 암호화 알고리즘, 비밀키 등 추가 정보의 관리 부실로 개인정보의 식별 내지 식별가능성이 발생할 우려가 있고, 암호화, 복호화의 과정에 따라서는 개인정보의 처리에 따른 활용이 제한되거나 불가능하게 될 여지가 있으므로 적정한 암호화 알고리즘을 구현하고 이를 사용하는 것이 중요하다.

덧붙여 개인정보를 가명처리하는 목적이 단순히 안전한 보관이 아니라 개인정보를 여러 가지 형태로 처리하여 개인정보처리자 등 개인정보이용 주체가 새로운 정보를 생산하거나, 정보주체에게 유익한 용역을 제공하는 것이므로 정보를 정확하고도 효율적으로 처리하는 과정에서도 암호화조치가 기여를 할 것이다.

아래에서는 항을 바꾸어 여러 가지 암호화조치의 유형과 그 실례를 간단히 살펴보겠다.

다. 암호화조치의 유형

1) 고전암호

암호는 기원 이전부터 정치적, 군사적 목적으로 다양한 형태로 사용되었는데, 로마의 통치자인 카이사르 암호, 아핀 암호, 아트바스 암호 등이 있으나, 이는 알파벳 등 평문의 문자배열을 일정한 규칙에 따라 재배열하여 암호화하

---

computation), 암호화폐(cryptocurrency), 블록체인(blockchain), 디지털 포렌식(digital forensic)등을 들 수 있다. 이상 김명환, 앞의 책 p.12-14

133) 김명환, 앞의 책, p.6-7

는 원시적인 형태의 암호화 알고리즘[134]으로 앞서 살펴본 것과 같이 각종 공격에 취약하여 이와 같은 암호화로는 신법의 암호화조치를 하였다고 볼 수 없다. 문제는 가명처리 가이드라인이 이러한 단순한 치환 형태의 암호화 알고리즘을 가명처리의 한 예로 들고 있다는 것이다.

2) 기계식 암호

제2차 세계대전 당시 독일군이 사용하던 암호기계 에니그마(enigma)는 회전판(rotor)의 시작위치가 비밀키이고, 회전판이 한번 돌때마다 26개의 알파벳이 뒤섞이면서 위와 같은 치환 암호를 반복 시행하게 되는데 그 과정에서 평문을 타자하면 암호문을 구성하는 알파벳이 램프에 전시되어 이로 암호문을 작성하며, 비밀키는 주기적으로 교체되었다고 한다.[135]

이미 널리 알려져 있다시피 위 암호는 치환 암호에 불과하였으나 당시 기술력으로는 단시간 내에 그 해독이 쉽지 않았다고 한다. 다만, 앨런 튜링(Alan Turing)이 현재의 컴퓨터와 같은 원리로 연산하는 기계를 개발, 이용하고, 예상 가능한 평문 중 독일군의 전문 서두에 반복되는 문구('Heil, Hitler')에 착안, 해독에 성공하여 전쟁수행에 도움을 주게 되었다.

이후 암호화 알고리즘은 컴퓨터의 연산능력의 발달과 함께 방대한 계산량이나 수학적 난제를 이용하는 방법으로 진화하여 현대에 이르고 있다. 다만 이와 같은 소위 기계식 암호는 현대 컴퓨터의 연산능력 범위 내에 있으므로 적정한 암호화조치로 사용할 수는 없다.

3) 비밀키, 대칭(symmetric)키 암호

대칭키암호는 암・복호화에 같은 키를 사용하는 알고리즘으로 이 경우 키를

134) 예를 들어 카이사르 암호의 경우 알파벳을 그 다음 세 번째로 암호화 하는 것(a→D, u→X)이고(김명환, 앞의 책, p.21), 아트바스 암호는 알파벳을 무작위의 알파벳 순서에 따라 대치하는 것(김명환, 앞의 책, p.29)이나 이러한 방식의 암호화는 복잡한 연산이 필요 없이 적당한 수의 암호문, 평문의 조합으로도 해독이 가능하다.

135) 이상 김명환, 앞의 책 p.44-45

비밀로 관리하여야 하기 때문에 비밀키암호라고 불리기도 하는데, 만일 여러 사용자 간에 동일한 키를 사용한다면 사전에 이를 공유할 수밖에 없다.

블록암호는 평문을 적당한 크기의 균등한 블록으로 나누어 블록 단위로 암호화하는 것으로 앞서 언급한 고전 암호 중에도 블록암호의 형태를 취한 것이 있으나, 현대의 블록암호는 컴퓨터 연산을 전제하고 있어 평문이 이진법의 비트로 전환되고 위 비트를 블록으로 나누어 암호화하게 되는데 단순히 대치(substitution)와 치환(permutation)을 반복하여도 안전성이 높아진다는 클로드 새넌(Claude Shannon)의 이론에 바탕을 두고 있다.[136]

암호화 알고리즘 중 하나인 AES(Advanced Encryption Standard)[137]는 대칭키암호로서 블록암호의 형태를 띠고 있는데, 그 안전성과 효율성을 평가받아 현재 상용되고 있는 것이다. 위 알고리즘은 평문을 128비트, 192비트, 256비트 등 일정한 블록으로 나눈 후, 바이트치환, 행이동, 열혼합, 같은 비트 라운드키의 혼합(비밀키로부터 생성) 등의 과정을 10~14라운드를 반복하여 암호화를 거치게 된다.

이와 같은 암호화 알고리즘이 공개되어 있지만 암·복호화에 사용되는 키(비밀키)가 송·수신자 또는 개인정보처리자에 의하여 비밀로 보관되어 있으므로 수학적인 공격, 즉 컴퓨터 연산으로 키를 알아내는 방법으로 해독하려면 키를 전수조사하거나 암호해독 알고리즘을 고안하여야 한다. 256비트의 비밀키를 사용한 AES 알고리즘의 경우, 상정 가능한 비밀키의 수가 2^{256}개에 달하므로 합리적인 시간 내에 해독하기 어렵다고 보고 있다.

AES는 중요한 정보를 안전하게 저장, 보관하면서도 효율적인 연산으로 처리할 수 있는 반면, 암호화된 정보를 사용하려면 비밀키를 이용하여 복호화하는 작업을 거쳐야 하므로 비밀키 또는 비밀키의 사용에 보안성이 요구된다.

이와 같은 AES 암호화 알고리즘은 공공기관, 금융 등 사회 전 분야에 널리

136) 김명환, 앞의 책, p. 146

137) 1977년 당시 미국 국가표준국(National Bureau of Standasrds)에서 공모 정차를 거쳐 미국의 공식 표준 암호체계로 채택한 것이 DES(data encryption standard)인데 여러 가지 공격 방법이 확인되어 국가표준국의 후신인 국가표준기술연구소(National Institute of Standard and Technology)에서 새로이 공모, 채택한 암호화 알고리즘이다. (김명환, 앞의 책, p.150, 175)

사용되고 있는데 애플(Apple)사에서 생산하는 휴대전화 아이폰에 내장된 iOS 운영체제에서의 AES 알고리즘 구현 방식을 아래와 같이 살펴본다.

iOS 내에서 AES 암호화 알고리즘을 작동시키는 AES엔진은 저장장치나 각종 프로세서로부터 격리되어 운영되는데[138], 이는 파일 암호화를 목적으로 고안된 프로그램으로서 휴대전화 기기의 생산과정에서 개별 기기마다 부여된 UID(「device's Unique ID key」의 약어로 기기마다 AES엔진에 사용될 목적으로 부여되는 키다.)에서 유도된 비밀키를 사용하여 AES 알고리즘을 가동하게 되고, 위 비밀키 역시 각종 저장장치와 프로세서로부터 물리적, 논리적으로 분리되어 있다고 한다.[139]

한편 파일 내 저장되는 개인정보 등 데이터의 보호에 대하여는, ① 휴대전화 이용 시 생성되는 파일마다 256비트의 랜덤한 '파일별 키'가 생성되고, 파일별 키는 위 AES엔진에 의하여 암호화되어 파일시스템에 저장된 파일의 메타데이터에 기록되며, ② iOS의 최초 설치 또는 기기의 초기화에 따라 위 UID로부터 유도되어 생성되는 키인 '파일시스템 키'로 위 메타데이터를 재차 암호화하는데, 파일시스템 키도 프로세서로부터 분리 관리되나, 사용자가 초기화 하거나 원격으로 명령하면 삭제할 수는 있고, 위와 같은 과정을 거쳤으므로 파일시스템 키를 삭제하면 알고리즘의 특성상 복호화가 불가능하다.[140]

위와 같은 iOS의 암호화 방식은 연산이 빠르고 비밀키의 전수조사가 어렵다는 AES의 장점을 이용하면서, 물리적, 논리적 분리 및 중층의 암호화 등의 방식으로 비밀키의 보호에 노력하여 수리암호가 지니는 위험성을 낮추려고 노력한 것으로 보인다.

4) 공개키, 비대칭(asymmetric)키 암호

---

138) Apple Inc., Apple Platform Security Guide(2022. 5.), p.9

139) 즉, 저장장치와는 별개의 칩(chip)에 기록되어 있다. Apple Inc, 앞의 자료, p.13

140) Apple Inc, 앞의 자료, p.73-74, 흔히 아이폰 기기의 잠금 암호의 문제가 회자되는데, 잠금 암호는 위 과정에서 사용되는 비밀키에 연동되는 것으로 보이고, 애플사의 설명에 따르면 6자리의 알파벳, 숫자 조합의 전수 공격은 5년 6개월이 소요된다고 한다. (Apple Inc, 앞의 자료, p.71)

비밀키 암호화 알고리즘을 취하는 경우 정보의 주고받는 송신자와 수신자가 상호 동일한 비밀키를 공유하여야 하는데 이는 보안성이나 효율성의 측면에서 문제점이 발생한다. 즉, 여러 사람이 동일한 비밀키를 공유하는 경우 비밀키의 보안이 취약해지고, 이를 보완하고자 송수신자의 쌍마다 별개의 비밀키를 생성, 공유하게 되면 다수의 송신자와 정보를 수수하는 수신자의 입장에서는 송신자 마다 비밀키를 구별하여야 하므로 효율적이지도 않고 그 과정에서 식별 가능성이 발생할 우려가 있다.

이에 대하여 1970년대 후반 등장한 공개키, 비대칭키 암호화 알고리즘은 암호화키를 공개함으로써 송수신자 간에 비밀키를 사전에 공유하지 않아도 암호문의 형식으로 정보를 주고받을 수 있는 알고리즘이다.

그 중 대표적인 예가 RSA 알고리즘으로 매우 큰 소수(150~200자리)의 소인수 분해가 어렵다는 수학적 난제를 기반으로 큰 소수 2개를 정하고 그로부터 계산해 낸[141] 암호화키를 공개하고, 복호화키는 수신자만 비밀로 간직한 후, 암호화키를 법으로 하여 계산(modulus)하는 방법으로 평문을 암호화하여 전송하면 이를 수신자가 자신만의 복호화키로 이를 복호화하여 정보를 확인하는데, 이 때 암호화키가 공개되어 있어도 암호문이 법계산의 결과이므로 이를 토대로 평문을 파악하는 것이 어렵기 때문이다.

RSA 알고리즘이 응용된 대표적인 예가 인증서를 통한 사용자의 전자서명을 통한 인증이다. 이는 RSA 알고리즘을 역으로 적용하여 송신자(인증을 받고자 하는 사람)가 인증서의 비밀번호를 입력하면 송신자만이 가지고 있는 비밀키가 활성화 되어 비밀키, 송신자가 공개한 키 및 법계산을 이용하여 평문에 해당하는 서명을 암호화 하여 전송하고, 송신자의 공개키를 가진 수신자가 전송된 암호문과 공개키를 이용한 법계산 결과를 이미 가지고 있는 인증정보와 비교하여 서명을 검증하게 된다.[142]

RSA 알고리즘은 상당한 연산이 필요하므로 방대한 정보의 양을 암호화하기에는 부적절[143]하지만 인증 등 중요한 기밀정보를 암호화하여 송신, 보관하는

141) 정확히는 두 소수(p, q)의 곱(pq)이 공개되는 암호화키(N)이고, (p-1)(q-1)과 서로 소인 정수 e(주로 65,537)는 공개하고, e와 d의 곱을 (p-1)(q-1)로 나눈 나머지가 1이 되는 d는 비밀키로 수신자만 가진다. (이상 김명환, 앞의 책, p.190 이하)

142) 이상 김명환, 앞의 책, p.239

것에 쓰일 수 있고, 비밀키를 사전에 공유하지 않아도 되는 장점이 있으며, 수학적 난제를 이용하였기에 외부 공격자가 암호문을 입수하더라도 컴퓨터 연산으로는 합리적인 시간, 비용으로 해독이 불가능하다.[144]

5) 해시함수를 이용한 암호화

해시함수는 입력된 임의의 길이의 메시지를 정해진 길이로 짧게 줄여 출력하는 함수[145]로, 아무리 용량이 큰 디지털데이터도 시간이 소요될 뿐 최후의 함숫값(이하 '해시값')은 함수가 예정한 비트길이로 산출될 뿐이다.

해시함수의 특징은 해시값을 알아도 원래의 입력값을 찾는 것은 거의 불가능하고(일방향성, onewayness), 같은 해시값을 가지는 두 입력값을 찾는 것도 거의 불가능하다는(충돌회피성, collision-freeness) 것이다.[146] 따라서 해시함수의 복호화는 수학적으로는 매우 어려운 것이어서 암호화만 가능한 알고리즘, 즉 일방향 암호화의 대표적인 예라고 하겠다.

이와 같은 특성에 따라, 주민등록번호, 생년월일, 성명과 같은 개인정보나 비밀번호와 같은 정보를 해시함수로 암호화하는 경우가 다수 있는데, 우선 해시값만으로 원래의 입력을 추측할 수 없는 점, 해시값이 다른 경우에는 원래 입력도 다르므로 데이터 사이의 구분이 가능한 점, 해시함수를 사용하는 방법이 간편한 점 등에 주목한 것으로 보인다. 실제 개인정보보호법과 동법 시행령은 비밀번호 등에 대하여 '일방향 암호화'를 의무화하고 있기도 하다.

또한 디지털데이터의 무결성, 동일성을 확인하고자 하는 경우, 디지털데이터를 구성하고 있는 모든 비트를 비교할 수 있을 것이나 이는 디지털데이터의 방대함으로 인하여 사실상 불가능한데, 비교 대상 디지털데이터의 해시값을 비교하면 무결성, 동일성을 쉽게 파악할 수 있다.[147]

---

143) 그런 이유로 방대한 데이터는 AES 알고리즘으로 암호화하고, 해당 비밀키를 RSA 알고리즘으로 암호화하여 전송하는 방법이 사용되기도 한다.(김명환, 앞의 책 p.195)

144) 다만, 아직 실용화가 되지 않은 양자 컴퓨터의 연산 방식으로는 인수분해를 의미 있는 시간 내에 상당한 확률로 해독 가능한 Shor알고리즘이 발표된 바 있어, 양자 연산이 실용화되면 안전성이 담보되지 않을 여지는 있다.(김명환, 앞의 책, p.188)

145) 김명환, 앞의 책, p.218

146) 김명환, 앞의 책, p.219

특히 기록이 담긴 블록을 구현하고 블록간의 연결, 기존의 블록을 새로운 기록이 담긴 블록에 연결하는 작업을 반복하게 되는 블록체인(blockchain) 기술의 개발, 구현에 있어서 해시값은 각 기록이 위, 변조되지 않았음을 입증하는 기본 전제가 되고 있다.[148]

이러한 해시함수는 MD4, MD5, SHA 등 다양한 함수로 개발되어 이용되어 왔는데, 연구자 등에 의하여 충돌쌍이 다수 발견되어 충돌 회피성에 의문이 제기된 함수는 사용이 배제되고 새로운 함수가 개발되고 있고[149], 심지어 구글과 같은 소위 '빅테크' 업체가 SHA-1의 충돌쌍을 직접 연구하면서 SHA-1을 사용하는 인증서는 지원을 종료하겠다고 발표하기까지 하였다.[150]

따라서 해시함수 연산과정에서 안정성이 확보될 수 있을 정도로 라운딩이 이뤄지는지, 해시값의 비트길이가 충분한지 등으로 안전성이 확보된 해시함수를 쓰도록 하는 것이 중요하다.

또한 비록 일방향 암호화이지만 원래의 데이터가 유한하고, 그 길이가 짧은 경우에는 사전에 암호문인 해시값의 목록을 '테이블'로 만들어두고 해시값을 비교하는 방식, 즉 전수조사 공격(brutal force attack)[151]이나 시간-메모리 교환공격(time-memory trade-off attack, TMTO)[152] 등으로 원래의 데이터를 알아볼 수 있는 위험성이 크다.

예를 들어 주민등록번호의 특성 상 가능한 평문은 누구든지 짐작할 수 있다. 즉 대략 146억 개[153]의 상정 가능한 주민등록번호가 있고, 146억 개의 주민등록번호를 모두 256비트 출력의 SHA-2 함수로 출력하여 사전에 테이블을

147) 비교 대상 데이터가 1비트만 달라도 해시값은 완전히 다르게 산출된다.
148) 김명환, 앞의 책, p.247-248
149) 김명환, 앞의 책, p.221
150) Google Removing SHA-1 Support in Chrome 56, Threat Post(2016, 11, 16,자 기사)
151) 가능한 비밀키를 모두 확인해 보는 방법(김명환, 앞의 책, p.22)으로 디지털데이터의 경우, AES 알고리즘에서 언급한 것과 같이 n비트길이의 비밀키에 대하여는 2^n회의 공격을 하여야 한다.
152) 비밀키를 찾는데 필요한 연산 횟수와 저장 공간을 상황에 따라 조정하는 기법으로 저장 공간의 크기에 따라 소요시간이 증감되는 것을 이용하는 것으로 사전 계산을 전제로 한다. (김명환, 앞의 책, p.158)
153) 앞자리는 010101부터 991231까지 36,500개(100×365)이고, 뒷자리는 1000000부터 4999999까지 4,000,000개이므로 146억 개의 경우의 수가 도출되나, 주민등록번호 뒷자리의 생성규칙에 따라 경우의 수는 더 감소할 것이다.

만든다면 그 과정에서 소요되는 시간은 수분에 불과하다.[154] 이와 같이 제작한 사전 테이블과 확보한 해시값을 위와 같은 공격방식으로 비교하는 연산을 컴퓨터로 처리하는 경우 상당한 메모리가 소요될 수 있지만, 이는 주민등록번호 생성규칙, 공격 대상의 성별, 연령을 특정하는 등의 방식으로 시간이나 메모리를 아낄 수 있을 것이다.

이러한 위험성을 완화하기 위하여 부가정보를 주민등록번호와 같은 개인정보 데이터에 더하여 해시함수를 적용하면 해시값의 경우의 수가 늘어나게 되어 위와 같이 사전에 테이블을 만들 때 추가의 계산량을 강요할 수 있는데 이를 솔트(Salt)라고 하고, 솔트에 사용되는 비트가 n이라면 2^n배의 계산을 더 하여야 한다.[155] 그런데 앞서 지적한 것과 같이 개보위의 가이드라인에서는 사실상 필수적인 요소인 솔트를 특별한 의미를 부여하지 아니한 채 단순히 예시로만 들고 있을 뿐이고 MDC와 같이 가명처리와 직접적인 연관이 없는 내용과 함께 소개하고 있을 뿐이다.

해시함수는 그 이용이 간편하고, 고유한 특성이 있어 널리 사용되고 있지만, 그만큼 허점에 대한 연구나 공격시도가 잦은 편이므로 안전한 해시함수를 이용하고, 민감정보 등 중요한 개인정보의 해시값의 관리는 더욱 엄격히 이뤄져야 할 것이다.

6) 동형암호(homomorphic encryption)

동형암호는 암호화된 데이터에 사칙 연산 등을 수행한 결과를 복호화한 값이 원 데이터에 같은 연산 등을 수행한 결과와 동일한 암호이다.[156]

---

154) 이는 주민등록번호나 비밀번호와 같이 평문의 분량(비트 수)이 적기 때문에 쉽게 사전 계산을 할 수 있을 때 발생하는 문제점이다. 주민등록번호를 이진수로 바꾼 것을 56비트, 7byte의 길이의 평문이라고 가정하고, 해시함수 SHA-256의 속도를 230Mib/s[인텔 코어i7-3770, 3.4GHz 프로세서로 처리한 속도를 제시한 결과가 있다.(cryptograms -Sha256 : Fast, pure and practical SHA-256 implementation, hackage.haskell.org, 2022. 5. 15. 확인)]로 가정하면, 146억 개의 주민등록번호(7byte 크기)의 평문 크기는 총 97465Mib에 해당하므로 이상과 같은 조건이라면 약 423초 정도면 테이블을 완성할 수 있게 된다.

155) Fenton, James L.; Grassi, Paul A.; Garcia, Michael E. (June 2017). "NIST Special Publication 800-63-3", p.54. NIST Technical Series Publications.

즉, 동형암호를 이용하면 개인정보를 암호화한 상태로 처리함으로써 정부, 사기업과 같은 개인정보처리자 등 제3자가 개인정보를 복호화하지 않아도 되는 효과가 있어 정보처리 과정에서는 물론, 정보이용 주체의 정보처리 중 내부자에 의한 개인정보침해 가능성까지 배제할 수 있어 유용한 암호화 알고리즘[157]이라 하겠다. 동형암호의 관건은 암호화한 상태에서의 연산속도인데, 연산속도가 계속 빨라지고 있는 추세라고 한다.[158]

동형암호 알고리즘이 실제 구현된 사례 중 정보주체의 동선을 코로나19 확진자의 동선과 비교하여 접촉 여부를 알려주는 어플리케이션의 사례[159]가 있다. 위 어플리케이션을 설치하면 사용자의 GPS 기반 위치정보를 10분 단위로 기록한 후, 사용자의 선택에 따라 접촉 위험 확인을 시도할 때 위치정보는 사용자의 휴대전화기기 내에 저장된 고유의 암호화키로 동형암호화 하여 서버에 전송되고, 서버에서는 암호화되어 전송된 정보와 익명화된 확진자의 동선과 비교하여 중첩 여부만을 계산한 후 사용자의 휴대전화기기로 되돌려 주며, 사용자의 휴대전화기기에서 이를 비밀키로 복호화하여 동선 중복 여부를 알려주는 방식을 취하게 된다.

이에 따라 정보처리를 하는 서버에서는 암호화된 사용자의 동선과 확진자의 동선의 중복 여부만 연산한 후 이를 사용자에게 돌려줄 뿐, 정보주체인 사용자의 인적사항, 동선을 알지 못하는 가운데 위와 같은 서비스를 제공하게 되고, 정보주체인 사용자의 선택에 따라 정보주체의 기기에서만 중복 여부를 확인할 수 있게 된다.

---

156) 김명환, 앞의 책, p.12, Cheon, Jung Hee; Kim, Andrey; Kim, Miran; Song, Yongsoo. Homomorphic encryption for arithmetic of approximate numbers, p.1 (2017)

157) 동형암호의 가능성을 확인하는 예로는 유전체와 같은 유전정보에 관한 연구에서 찾을 수 있다고 한다. 유전체의 경우 환자의 성명과 번호 이외에도 유전체 고유의 다형성에 따라 유전체 형태 자체로 식별가능성이 있으므로 염기 서열의 형태로 연구 자료를 제공할 것이 아니라, 암호화한 이후 암호문에 대한 분석만 허용하는 방식을 도입하여야 한다는 주장이 그것이다.(이원복, 앞의 논문, p.213)

158) Kristin Lauter; Michael Naehrig; Vinod Vaikuntanatha, Can Homomorphic Encryption be Practical?, p.114, CCSW '11: Proceedings of the 3rd ACM workshop on Cloud computing security workshop (October 2011)

159) 이하 Jung Hee Cheon and 6 others, Privacy Preserving COVID-19 Contact Tracing with Homomorphic Encryption. International Conference on Appropriate Technology(ICAT) 2021

라. 정리

이상 검토한 것과 같이 암호화조치에는 다양한 암호화 알고리즘이 있고, 사전적 의미의 암호화를 구현할 수 있는 알고리즘은 매우 다양하다.

개인정보의 보호 측면에서는 안전성이 전혀 없다고 보아야 할 부류에서부터 안전성이 충분하나 구현이 쉽지 않은 것들이 있고, 개인정보의 이용 측면에서도 처리 대상 정보에 따라 활용가능성이 천차만별이라고 하겠다.

마찬가지로 개인정보의 보호 측면에서만 접근한다면 개인정보를 완전히 삭제하거나, 익명정보의 형태로 개인정보를 처리하면 좋을 것이나, 그러한 경우 개인정보의 이용 가능성이나 활용 가치가 크게 떨어지는 것도 불가피하다.

이하에서는 이와 같은 암호화조치에 관한 내용을 염두에 두고 기존에 하급심 판례가 다루었던 비식별조치의 적정성 관련 사례를 검토한 후, 이를 토대로 신법의 해석, 신법에 대한 보완 논의에 있어 암호화 조치가 어떠한 의미를 지니는지 검토해 보겠다.

3. 기존 판례의 해석론

신법이 개정, 시행된 이후 개보위 단계에서 유권해석을 한 사례는 다수 있으나, 법원이 신법을 적용하여 판단한 민형사상 판결례는 찾아보기 힘들다. 다만, 구법 시행 당시에도 식별가능성 여부 또는 비식별조치의 적정성을 판단하여 개인정보보호법이 보호하는 개인정보인지 여부를 판단한 사례가 확인된다.

따라서 아래에서는 그와 같은 사안의 사실관계, 당사자의 주장, 법원의 판단을 간략히 소개하면서 이를 통하여 이상 살펴본 신법에 대한 해석론, 외국 입법례를 토대로 앞으로 신법을 어떻게 해석, 적용하고 또 보완해 나갈 수 있을지 고민해 보고자 한다.

가. 식별가능성의 판단 주체 : 식별주체

어떠한 정보가 개인정보인지에 관한 논의는 본 연구의 직접적인 주제는 아니나 법에서 보호하는 개인정보는 '살아 있는 개인의 정보'로서, 첫째, 성명, 주민등록번호 및 영상 등을 통하여 개인을 알아볼 수 있는 정보이고, 둘째, 해당 정보만으로는 특정 개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 알아볼 수 있는 정보(이상 법 제2조 제1호)이므로 결국 정보 자체로 정보주체를 알아볼 수 있는지, 즉 식별가능성이 있는지에 따라 달려 있음은 이미 수차례 정리한 것과 같다.

문제는 신법이 도입한 가명처리, 익명정보의 개념에서 적정한 가명처리가 되었는지, 더 이상 정보주체를 알아볼 수 있는지 여부에 대한 판단에서도 식별가능성이 있는지 여부가 기준이 될 수밖에 없으므로[160], 법에 따라 보호되는 개인정보인지 여부를 식별가능성의 유무에 따라 판단한 기존의 사례에서 우선 식별가능성의 판단 기준 중, 판단의 주체 내지 시각, 즉 식별주체에 관한 사례를 간략히 검토할 필요가 있다.[161]

1) 하급심 사례

가) 휴대전화의 IMEI[162]와 USIM[163]의 일련번호 수집 사건

160) GDPR, HIPAA의 예에서 '식별 가능한'으로 번역할 수 있는 'identifiable'이 반복되는 것도 같은 이유로 볼 수 있다. 즉, 식별되지 않고, 더 이상의 식별가능성도 없다면 익명정보가 될 것이고, 현재로서는 식별하기 어려우나, 추가 정보가 결합되면 식별될 수 있다면 가명정보가 될 것인데 위와 같이 처리하는 것 자체가 결국 식별가능성을 없애는 과정이고, 그 결과 또한 식별가능성의 유무로 판단할 수밖에 없다.

161) 이와 관련하여 대법원에서 명시적으로 논한 사례는 확인되지 않는다. (고학수 등 7명, 앞의 책, p.19)

162) 국제 휴대전화 식별번호(International Mobile Equipment Identity)의 영문 약자로, 제조사가 단말기를 제작할 때 기기마다 고유한 번호로 부여하는 것이다. [https://www.techopedia.com/definition/5066/international-mobile-equipment-identity-imei(2023. 1. 1. 확인)]

163) 범용 가입자 인증 모듈(Universal Subscriber Identity Module)의 영문 약자로 휴대전화에 장착하는 IC회로 칩 형태로 정보를 담는 것인데, 통상 인증 관련 데이터가 저장된다. 이상

위 사건은 어플리케이션 개발자들이 증권시세 검색 어플리케이션을 개발하면서 사용자가 자신의 관심종목의 시세 등 거래 정보를 신속히 검색, 확인할 수 있도록 하기 위하여 사용자들이 어플리케이션을 설치하면 어플리케이션이 사용자로부터 동의를 받지 아니한 채 사용자의 휴대전화 기기의 IMEI와 USIM의 일련번호, 또는 IMEI와 휴대전화번호를 수집하도록 한 사안이다.

재판부는 "종전에 비하여 정보들이 쉽게 결합되어 개인을 식별할 수 있게 되었고 기계적인 정보라 하더라도 특정 개인에게 부여되었음이 명백하고 이러한 정보를 통하여 개인이 식별될 가능성이 있다면 개인정보로 봄이 상당하다."고 전제하면서, 위 각 정보 모두 그 자체로는 사용자정보를 확인할 수는 없음을 인정한 후, "IMEI는 각 휴대전화 기기에 부여된 고유번호이고, 사용가능한 IMEI의 목록은 통신사에서 관리하며, 권한 있는 자가 여러 가지 관련 정보를 조합하면 사용자 정보를 확인할 수 있는 점, USIM의 일련번호도 IMEI와 마찬가지로 통신사에서 가입자의 인적사항과 함께 관리하므로 권한 있는 자가 여러 가지 관련 정보를 조합하면 이를 확인할 수 있는 점, 이 사안에서 피고인들도 동일성 인증을 하기 위한 ID 대용으로 각 정보를 수집한 점, 정보주체가 작성하는 휴대전화 가입신청서에 IMEI, USIM의 일련번호가 기재되고 이들이 위와 같이 통신사에서 데이터베이스의 형태로 관리하는 점 등"을 종합하면, 피고인들 외에 통신사 등 제3자들이 가지고 있는 정보와 결합하는 경우에는 단순한 휴대전화 기기나 IC칩이 부착된 카드의 고유번호에서 나아가, 특정 정보주체가 소유하는 휴대전화의 기기번호, USIM의 일련번호라는 의미를 알 수 있게 된다고 보아[164], 식별주체를 개인정보처리자로 한정하지 않고 통신사 등 제3자로 넓혀 판단한 후, 개인정보성을 긍정하였다.

나) 휴대전화번호

---

https://www.igi-global.com/dictionary/usim-universal-subscriber-identity-module/31282 (2022. 11. 1. 확인)

164) 서울중앙지방법원 2011. 2. 23. 선고 2010고단5343호 판결

한편 피고인이 대출신청자들의 휴대전화번호 15만 개를 포함한 다수의 개인정보를 임의로 타에 양도하였다고 공소제기된 사안에서 재판부는 통신사에 의하여 하나의 휴대전화번호가 한 명의 사용자에게 부여되는 점, 통신사가 가입 약정 시에 제공받은 성명, 주민등록번호 등 개인정보를 관리하고 있는 점, 사용자가 휴대전화를 사용하는 기간 동안 당해 휴대전화번호는 위 사용자에 전속하는 점 등을 종합하면, 통신사가 관리하는 정보와 쉽게 결합하여 정보주체를 알아볼 수 있으므로 개인정보에 해당한다고 보았다.[165]

심지어 정보주체의 휴대전화번호 뒷부분 4자리도 개인정보로 인정한 사례가 있는데, 당해 사안에서 재판부는 정보주체가 휴대전화번호 뒤 4자리를 선택할 수 있고 어떠한 의미나 패턴(생일 등 기념일, 한 가족의 동일한 뒷부분 사용, 일반 유선전화번호의 뒷부분과의 일치, 영업용 번호의 특징 등)을 담기도 하므로, 정보주체와 인적 관계가 있다면 4자리로도 정보주체를 알아볼 수 있다고 보아 개인정보로 보기까지 하였다.[166]

이상 사례는 통신사 외에도 개인정보처리자 등의 사정을 폭넓게 고려하여 개인정보의 인정 범위를 넓게 해석한 예로 볼 수 있다.

다) 혈액검체용기에 기재된 검체번호, 검사항목, 검사결과 수치 등

그런데 의료기관의 직원이 채혈된 혈액의 검사결과가 기재된 혈액검체용기를 임의로 반출하였다는 내용으로 개인정보보호법위반으로 공소제기된 사안에서 재판부는, “단순히 정보제공자를 기준으로 판단할 것이 아니라 정보의 내용, 정보 수수자 간의 관계, 정보 수수의 목적, 방법, 결합을 위하여 필요한 노력, 비용 정도 등을 합리적으로 고려하여 (개인정보 해당 여부를) 판단하여야 한다.”고 보면서도, 「검체용기에 기재된 내용 자체는 개인정보로 보기 어려운 점, 그 내용을 토대로 개인정보를 식별하려면 의료기관 내의 시스템을 이용하여야 하는데 그 접근 권한이 직책에 따라 다른 점, 피고인들이 개인정보로 볼

165) 서울중앙지방법원 2015. 1. 14. 선고 2014고단5061호 판결(제1심), 같은 법원 2015. 4. 9. 선고 2015노387호 판결(제2심)
166) 대전지방법원 논산지원 2013. 8. 9. 선고 2013고단17호 판결

수 있는 내용을 삭제하고 반출한 사정을 고려할 때 피고인들에 대하여 범의를 인정하기도 어려운 점」 등을 들어 범죄의 증명이 부족하다고 보았다.[167]

2) 검토

이상 살펴본 것과 같이 일반적으로 법원은 식별가능성의 판단 기준, 시점 내지 시각을 최소한 개인정보처리자에 국한하지는 않는 것으로 보이고, 이는 구법에서부터 현재까지 유지되고 있는 법률상 문언인 「다른 정보와 쉽게 결합하여 알아볼 수 있는 경우」의 해석에 충실한 것으로 보인다.

다만, 위 혈액검체용기 관련 사안에서는 여러 가지 사정을 고려하여야 한다고 하면서도 개인정보처리자에 관한 사정, 심지어 그들의 범의가 인정되기 어렵다는 점까지 들어 개인정보에 해당하지 않는다는 판단을 하기도 하였다.

이와 관련하여 강학상 식별가능성의 판단을 어떠한 주체의 시각에서 보아야 하는지에 대하여, 소위 절대설, 상대설의 구분으로 설명하는 견해가 다수 있다.[168] 이와 같은 논의가 펼쳐지는 이유는 식별가능성을 판단할 때 식별주체가 누구인지에 따라 그 결론이 달라질 여지가 크고, 실제 아래에서 계속 살펴보게 되는 법원의 판단에서 영향을 끼치고 있기 때문이다.

절대설에 따르면 모든 가능성을 염두에 두고 식별주체가 누구든 간에 식별가능성이 있다면 식별가능성을 긍정하는 것이고, 상대설은 해당 정보처리자의 관점, 즉 식별주체를 정보처리자에 국한하는 것이라고 한다.

절대설에 의할 경우 잠재적 식별주체가 존재할 수 있는 사소한 가능성까지 고려하는 바람에 식별 가능하다고 인정하는 범위가 넓게 되어 필요 이상으로 강력한 비식별조치를 요구할 우려가 있고, 반면에 상대설에 의할 경우 개인정보의 보호범위가 극단적으로 좁혀지게 됨은 자명하다.

따라서 개인정보처리자 이외의 제3자도 식별주체에 포함시키되, 당해 정보에 대하여 식별시도를 할 수 있는 합리적 가능성이 있는 사람들까지 식별주체

---

167) 수원지방법원 성남지원 2017. 9. 15. 선고 2017고단1438호 판결(제1심), 수원지방법원 2018. 4. 12. 선고 2017노7275호 판결(제2심)

168) 양기진, 앞의 논문, p.62 이하 및 고학수 등 7명, 앞의 책, p.19 이하

로 제한하는 것이 상당하다.[169]

구체적으로는 비밀번호, 인증정보나 법령상 특별한 취급을 받는 고유식별정보, 민감정보와 같이 정보주체에 심각한 영향을 끼칠 수 있는 정보의 경우, '해커'와 같은 잠재적 공격자인 제3자의 시간, 능력, 정보까지 포함하여 식별가능성을 판단할 필요가 있고, 휴대전화번호와 같은 정보의 경우, 정보의 특성을 고려하여 통신사와 같이 관련 정보를 보유하거나 관련 정보와 관련되는 사람까지 포함시킬 수 있을 것이다.

덧붙여 이와 같은 식별주체에 관한 내용은 결국 규범적 판단의 기준이므로 일종의 법률 요건에 해당할 것인데, 위와 같이 실제 법적 판단에 중요한 영향을 끼치고 있으므로 법령의 단계에서 규율될 필요가 있다.

요컨대, 식별가능성의 판단에 있어서는 개인정보처리자의 관점에서만 판단하여서는 아니 되고 정보의 내용, 이용 환경, 관리 실태, 취득경위, 개인정보처리자의 사정 등을 종합하여 합리적 범위 내의 제3자의 관점까지 고려하여야 함이 상당한데 앞에서 정리한 것과 같이 GDPR은 전문 제26조에서 「해당 정보처리자 또는 다른 자(제3자)에 의하여 사용될 합리적인 가능성이 있는 모든 수단」을 고려하여 식별가능성을 판단하여야 한다고 규정하여 입법으로서 명확히 규율하고 있으므로 우리에게 시사하는 바가 크다.[170]

나. 비식별조치의 적정성

1) 2012~2013년 카드 3사 정보유출 사건

가) 사실관계[171]

---

169) 이와 유사한 입장을 절충설로 표현하기도 한다. 양기진, 앞의 논문, p.63

170) 이와 관련하여 온라인미디어서비스제공자가 인터넷서비스제공자로부터 법적 조치를 통하여 정보주체를 식별할 수 있는 정보를 얻을 수 있는 이상, 온라인미디어서비스제공자가 저장한 유동 IP 주소도 개인정보에 해당한다고 본 유럽사법재판소의 판결(Breyer v. Germany 사건)이 있었다고 한다. (양기진, 앞의 논문, p.66 이하)

171) 서울중앙지방법원 2016. 7. 15. 선고 2015고합336호 판결 및 항소심 판결인 서울고등법원 2020. 1. 31. 선고 2016노2150호 판결의 범죄사실 등을 정리하였다.(상고기각 확정)

각 피고인 회사는 신용카드 등 업무를 취급하는 회사이고 A회사는 금융기관들이 공동출자하여 설립한 회사로, 갑은 A회사 직원으로 각 회사에 파견되어 신용카드 부정사용탐지시스템의 개발업무를 담당한 사람이다.

이 때 갑은 2012. 6.부터 2013. 12.까지 수차례에 걸쳐 각 회사 사무실 내에서 A회사를 위하여 업무를 수행하면서 각 회사가 보유하고 있던 고객들의 개인정보(성명, 주민등록번호, 휴대전화번호, 주소 등) 1,700여만 건에서 4,321만여 건 가량을 자신의 USB 메모리 저장장치에 몰래 저장하여 반출하였다.

검사는 각 피고인 회사들에 대하여 각 회사의 개인정보보호책임자 등 사용인들이 구 개인정보보호법(2015. 7. 24. 법률 제13423호로 개정되기 전의 것) 제24조 제3항(현행법과 동일)에 규정된 고유식별정보에 대한 안전성 확보에 필요한 조치(공유폴더의 권한 미설정, 보안프로그램 미설치. USB 메모리 반입 허용, 고유식별정보의 암호화조치의무 불이행 등)를 취하지 아니하였다고 보아 양벌규정에 따라 각 회사들에 대하여 공소제기하였다.

나) 각 회사의 주장

각 회사들은 다양한 논점에서 공소사실을 부인하였는데, 그 중 본 연구와 관련된 부분인 '암호화조치 불이행'만 정리하여 보면, 구법 당시 행정안전부 고시로 규정되었던 개인정보의 안전성 확보조치 기준 제6조 제5항에 따라 '위험도 분석'을 거쳐 고유식별정보에 대하여 암호화하지 않기로 한 것이고, 당시 신용정보법이 주민등록번호 등의 암호화를 규정하지 않았으며, 암호화를 하였다고 하여도, 갑이 수행하는 작업의 특성상 A회사에 제공할 때에는 복호화를 하거나 복호화키를 주었을 것이므로 인과관계가 없다고 주장하였다.

다) 법원의 판단

제1,2심 각 재판부는 개인정보보호법 위반행위에 대한 고의가 있으면 족하고, 그 결과, 즉 분실, 유출, 도난 등의 결과까지 의욕할 필요가 없다는 전제를

설시한 후, 이 사건에서는 개인정보를 내부망에서 저장한 것이 아니라 외부의 컴퓨터 등에 옮긴 상태였으므로 암호화조치의무가 그대로 발생하는 점, 암호화를 하지 않은 이상 공유폴더, USB 메모리, 보안프로그램에 대한 관리, 운영을 철저히 하였어야 하는 점 등을 들어 각 피고인 회사들에 대한 죄책을 인정하였다.

라) 정리

위와 같은 법원의 결론은 문제된 개인정보의 보관, 관리 상황을 자연적으로 관찰한 결과에 따라 의무 발생, 준수여부를 합리적으로 판단한 것으로 평가할 수 있다.

특히 행위자에 대하여 개인정보보호법에서 부과한 안전성 확보조치를 위반하는 것에 관한 고의를 인정함에 있어서 위반행위로 인한 결과, 즉 개인정보의 분실, 유출, 도난, 훼손 등의 결과에 대한 고의까지 요구하지는 않겠다는 판단은 형사법의 기본 법리에 충실한 견해이자, 개인정보의 보호에 주목한 것이고, 특히 신법에서도 당연히 적용되어야 할 법리를 선언한 것이라고 생각된다. 첨언하자면 신법 제73조 제1호에는 법 제28조의4 제1항을 위반행위로 추가하여 개인정보에 대한 적정한 가명처리의무를 위반한 경우를 구성요건으로 삼고 있으므로 같은 법리를 적용할 수 있을 것이다.

다만, 위 사건에서 현행 개인정보의 안전성 확보조치 기준의 맹점이 드러나므로 유의하여야 한다. 즉, 구법 이래 현재까지 일정한 요건에 따라 정보통신서비스 제공자가 아닌 개인정보처리자에 대하여 고유식별정보에 대한 암호화조치의무를 부담하지 않도록 허용하고 있는데, 이 사건에서 알 수 있듯이 개인정보의 분실, 유출, 도난 등의 결과는 외부자뿐만 아니라 내부자(또는 개인정보처리자로부터 권한을 수여받은 자)에 의하여서도 일어날 수 있으므로 이는 시정될 필요가 있다.

요컨대 이와 같은 개인정보침해행위는 내부망에 보관하거나 개인정보처리자의 통제 하에서도 발생할 수 있는 점, 암호화조치를 취하게 되면 일단 암호화된 데이터가 분실, 유출, 도난이 되더라도 암호화키를 분리 보관하는 경우 개

인정보가 실제 식별될 가능성이 낮아지는 점, 위와 같은 안이한 고시의 규율로 인하여 이 사건 각 피고인 회사와 같이 고시의 규정을 소위 '면피용'으로 내세울 우려가 있는 점 등에서 최소한 개인식별정보와 같이 중요한 개인정보에 대하여는 반드시 암호화조치의무를 부과하여야 할 것이다.

2) 약학정보원 조제정보 유출 사건

가) 기본적 사실관계[172)]

약학정보원은 대한약사회가 주도하여 2001. 2.경 설립한 재단법인으로 국내 제조, 수입의약품 등에 대한 정보 수집, 데이터베이스 구축 등을 목적으로 하고 있고, 공교롭게도 IMS Health는 본 연구의 미국의 입법례 관련 부분에서 등장한 회사로, 사안에서 등장하는 H회사는 위 IMS Health가 100%의 지분을 가지고 설립한 한국 자회사로 IMS health를 위하여 의약 및 건강 관련 사업에서의 제품 및 서비스에 관한 자료수집, 시장조사 보고서 작성 등을 목적으로 하는 회사이다.

대한약사회는 약국관리 프로그램인 IMS2000을 개발하고 약사들에게 무료로 제공하여, 약사들이 위 프로그램을 이용하여 처방, 조제 데이터를 입력하고 이를 토대로 건강보험심사평가원에 요양급여를 청구하며 재고관리를 할 수 있도록 하였으며, 약학정보원은 위 프로그램의 관리, 운영을 대한약사회로부터 위탁받았다.

한편 약학정보원과 H회사는 2010. 12. 30. 데이터공급계약을 체결하고 약학정보원이 2011. 1. 1.부터 2015. 12. 31.까지 위 프로그램을 사용하는 약국으로부터 수집한 재고, 처방, 조제에 관한 데이터를 H회사에 제공하고, H회사는 그 대가로 매년 3억 원 및 H회사의 매출에 따른 가산금을 더한 금원을 지급하기로 하였다.

172) 서울중앙지방법원 2020. 2. 14. 선고 2015고합665호 등 판결, 서울고등법원 2021. 12. 23. 선고 2020노628호 판결, 서울중앙지방법원 2017. 9. 11. 선고 2014가합508066호 등 판결, 서울고등법원 2019. 5. 3. 선고 2017나2074963호 등 판결의 범죄사실, 인정사실 등을 요약하였고, 현재 각 사건은 상고심 계속 중이다.

이에 따라 약학정보원의 임직원들은 2011. 1.경 각 약국에서 위 프로그램에 저장한 처방전 관련 정보(환자 인적사항, 질병, 처방 의약품의 내역 등)를 약학정보원의 서버로 자동 전송하는 프로그램을 개발하여 이를 위 IMS2000의 업데이트 파일에 포함하고, 약사들로 하여금 설치하도록 하였는데, 그 과정에서 팝업(pop-up) 창으로 동의("약학정보원이 사용자가 IMS2000에 저장하는 조제정보를 통계작성 및 학술연구 등의 목적으로 수집할 수 있고, 조제정보는 탑재된 프로그램에 의하여 암호화하여 개인을 알아볼 수 없는 형태로 약학정보원에 제공된다."는 취지)를 구하면서 동의하여야 프로그램이 실행되게 하였다.

위와 같은 과정에 따라 약학정보원은 2011. 1.부터 2015. 1.까지 약 140억 개 가량의 조제정보 등을 H회사에 제공하고, H회사는 해당 정보를 위 미국 본사의 서버에 전송하여 미국 본사는 이를 분석하여 제약회사 등에 판매하였다.

나) 암호화조치 방식

이 사건이 학계는 물론 사회 일반에서 회자되고 있는 이유는 적정한 암호화 조치 여부가 하나의 쟁점이 되었기 때문인데, 약학정보원, H회사가 주민등록번호, 성명, 생년월일 등 개인정보를 암호화한 방식은 아래와 같다.

⑴ 1차 암호화 : 알파벳 치환에 의한 암호화

약학정보원은 2011. 1.부터 2014. 6.까지는 환자 주민등록번호 13자리 중, 홀수 자리와 짝수 자리의 각 숫자를 아래 표와 같은 규칙에 따라 영어 알파벳 소문자 13개로 치환한 다음, 가장 앞에는 치환한 알파벳의 13번째 문자를, 가장 뒤에는 치환한 알파벳의 6번째 문자를 소위 '노이즈(noise)'로서 추가하는 방식으로 암호화하였는데 이는 H회사 소속 직원이 고안하여 약학정보원 측에 제안한 것이었다.

- 치환 방식(형사 제1심 판결서에서 발췌) -

변환

홀수자리	1	2	3	4	5	6	7	8	9	0					
알파벳소문자	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
짝수자리						0	9	8	7	6	5	4	3	2	1

예) 541008-1030024 → elafjhafcfjnd

㉡ 2단계: 변환된 문자열의 6번째 자리의 문자를 문자열의 가장 마지막 자리에, 13번째 자리의 문자를 첫 번째 자리에 추가하여 15자리의 알파벳 문자열을 생성

예) 541008-1030024 → elafjhafcfjnd → delafjhafcfjndh

⑵ 2차 암호화 : SHA512(앞서 살펴본 SHA-2 계열의 해시함수로서 512비트의 해시값이 출력되는 것이다.)에 의한 암호화

이어 약학정보원은 2014. 6.부터 9.까지는 주민등록번호를, 2014. 10.부터 2015. 1.까지는 성명, 생년월일, 성별을 각 SHA-512로 암호화하였는데, 이는 행정안전부의 지적에 따라 암호화 방식을 변경한 것이라고 한다. 그런데 이때 약학정보원은 주민등록번호의 해시함수값과 알파벳치환 암호문을 대응시키거나, 주민등록번호의 해시함수값과 성명, 생년월일, 성별의 해시함수값을 대응시키는 방법으로 작성된 '매칭테이블'을 별도로 제작한 후, 이를 저장매체에 저장하여 H회사 측에 제공하였다.

덧붙여 검사가 2015. 1. H회사의 사무실을 압수수색한 결과 주민등록번호가 알파벳 치환되어 표시된 처방 및 조제정보 4,336,993,835건이 저장되어 있었고, 이를 복호화한 결과 48,331,005명의 주민등록번호가 확인되었다고 한다.

다) 약학정보원 등의 주장

⑴ 민사 : 불법행위에 기한 손해배상금 청구

의사, 환자 등 1,876명이 원고가 되어 약학정보원, H회사 등에 대하여 위와 같은 개인정보의 수집, 처리가 구 정보통신망법 제49조를 위반하여 타인의 비밀을 침해, 도용하는 한 경우, 같은 법 제49조의2에 따라 정보통신망을 통하여 속이는 행위로 정보를 수집한 경우 또는 정보주체의 동의 없이 개인정보(민감

정보)를 처리한 경우로서 불법행위에 해당한다고 주장하였다.

이에 대하여 약학정보원 등 피고 측은 첫째, 주민등록번호 등에 대하여 암호화 등 비식별화 조치를 거친 이상 '개인정보'에 해당한다고 볼 수 없고, 둘째, 이를 개인정보로 보더라도 개인정보보호법 제18조 제2항 제4호에 따라 통계분석을 위하여 제공된 것으로 적법하며, 셋째, 원고들에게 발생한 손해가 없다고 주장하였다.

⑵ 형사 : 개인정보보호법위반

검사는 약학정보원이 정보주체의 동의 없이 민감정보인 조제정보 200만 건을 위 프로그램을 이용하여 수집, 저장, 보유하고, 540만 건을 H회사에 제공한 부분, H회사가 위 540만 건을 미국 본사에 전송한 부분을 각 개인정보보호법 위반에 의율하여 공소제기하였는데, 알파벳 치환에 의한 방식은 H회사에서 제공한 것인 점, 복호화가 가능한 경우에는 여전히 개인정보로 보아야 하는 점, 약학정보원과 H회사 상호간에 암호화, 복호화 관련 자료를 공유한 점, 약국을 기망하여 업데이트를 한 점 등을 들어 약학정보원 등 관계자들에 대하여 개인정보보호법위반이 인정된다고 주장하였다.

이에 대하여 약학정보원 등 피고인 측은 암호화조치로 인하여 비식별화가 이루어졌으므로 개인정보, 민감정보에 해당하지 아니하는 정보이고, 법위반의 범의가 없었다고 보았다.

라) 법원의 판단

각 사건마다 민사의 경우 청구원인이 다양하고, 형사의 경우에도 쟁점이 여러 가지이나, 본 연구에 직접 관련된 부분에 대한 법원의 판단을 정리하여 보면 아래와 같다.

⑴ 민사 : 불법행위에 기한 손해배상금 청구

㈎ 제1심 : 원고 청구기각(서울중앙지방법원 2017. 9. 11. 선고 2014가합508066호 등)

제1심 재판부는 암호화 등 적절한 비식별화조치를 취함으로써 특정 개인을 식별할 수 없는 상태에 이른다면 이는 개인정보에 해당할 수 없을 것인데 이 사안에서는 알파벳 치환 방식에 의한 1차 암호화의 경우 암호화 규칙이 단순한 점, 수사기관에서도 쉽게 복호화한 점, H회사 측에서 암호화 방식을 먼저 제안한 점 등에서 재식별가능성이 현저하므로 암호화조치 이후에도 여전히 개인정보에 해당하나, SHA-512에 의한 2차 암호화의 경우 일방향 암호화로서 복호화가 원칙적으로 불가능한 점, 달리 재식별화할 경제적 유인이 없었던 것으로 보이는 점, 당사자 중 누구도 복호화를 시도하지는 않은 점 등에서 비식별조치가 이뤄진 것으로 봄이 상당하여 개인정보에 해당하지 않는다고 판단하였다.

계속하여 위 재판부는 1차 암호화의 경우, 적법한 개인정보의 처리로 볼 수는 없으나, 다른 분야에 유출되거나 범죄에 사용되지는 않은 점, 제3자의 입장에서는 암호문으로 개인을 식별하기 어려운 점, 제3자가 암호문에 접근할 가능성이 없는 점 등에서 정신적 손해가 실제 발생하였다고 보기 어렵고, 2차 암호화의 경우 개인정보보호법 제18조 제2항 제4호에 따라 통계작성을 위하여 허용된다고 보았다.

㈏ 제2심 : 원고 청구기각(서울고등법원 2019. 5. 3. 선고 2017나2074964호 등)

제2심 재판부는 결론은 제1심과 같았으나. 위와 같은 암호화조치는 모두 적정한 비식별조치로 볼 수 없다고 보았다.

즉, 위 재판부는 1차 암호화의 경우 제1심과 유사한 취지로 적정한 암호화조치로 볼 수 없다고 보았고, 2차 암호화의 경우 비록 어느 정도 안전성이 인정된 SHA-512 알고리즘을 사용하여 일방향 암호화조치를 취한 이상 그 자체로는 복호화 가능성이 낮았다고 하더라도, 약학정보원과 H회사 상호 간에 매칭테이블을 주고받은 이상 쉽게 복호화가 가능한 점, 특히 식별가능성을 판단함

에 있어 정보 수령자의 주관적 의도, 동기, 이익, 활용방법까지 고려할 수는 없는 점 등에서 적정한 비식별조치가 인정되지 아니하므로 어느 방식에 의하였든 간에 여전히 개인정보로 봄이 상당하다고 보았다.

또한 위 재판부는 여전히 개인정보로 볼 수 있는 정보를 정보주체의 동의 없이 수집, 처리한 이상 적법하게 수집한 정보를 전제한 개인정보보호법 제18조 제2항 제4호를 적용할 수 없으므로 약학정보원과 H회사는 개인정보보호법을 위반하였다고 잠정적으로 결론지었다.

다만, 위 재판부는 제3자에게 개인정보가 유출되지는 않은 점, H회사나 미국 본사가 처방정보를 매입한 목적과 다르게 이용하지는 않은 점, 제3자의 입장에서는 개인정보를 식별하기 어려웠을 것으로 보이는 점, 미국 본사의 서버에서도 이미 삭제된 점 등에서 정신적 손해가 실제로 발생하였다고 보기는 어렵고 판단하였다.[173]

⑵ 형사 : 해당부분은 무죄

각 형사 재판부에서는 1, 2심 공히 약학정보원 등의 관계자인 피고인들에 대하여 개인정보보호법위반의 고의가 있었음을 입증할 증거가 부족하다고 보아 해당 공소사실에 대하여는 무죄를 선고하였다.

우선 재판부는 재식별가능성, 복호화 가능성이 합리적으로 존재한다면 여전히 개인정보에는 해당한다고 보면서도, 행위자가 그 정보를 식별가능한 개인정보로 인식하였는지 여부가 인정되어야 한다고 전제한 뒤, 이 사안에서의 고의는 비식별화 또는 암호화된 개인정보를 식별화 또는 복호화하여 처리할 수 있다는 인식과 함께 식별 가능한 상태로 치환하여 처리하는 것을 용인하는 의사까지 확인되어야만 인정될 수 있다고 보아, 고의의 인정 여부에 필요한 사실관계를 넓히고 있어 위 신용카드사 사건에서의 재판부와는 견해를 달리 하고 있다.

---

173) 다만, 이 사건 이후인 2015. 7. 24. 개정되어 신설된 법 제39조 제3항, 제4항(법률 제13424호)에 따르면 정보주체인 원고가 손해배상액을 입증할 의무가 다소 경감되었다고 볼 수 있다.

이어 재판부는 사안에서, 약학정보원과 H회사는 상호 계약상 의무를 이행한다는 인식에 따라 암호화조치를 취한 점, 수집과정에서 다소 형식적이나마 약국 운영자들로부터 동의를 받은 점[174], 1차 암호화 방식을 취한 것은 일종에 과실에 불과한 점, 당사자들 간에 복호화할 이유나 동기가 없었던 것으로 보이는 점, 개인정보보호법이 위반행위 도중 시행된 점, 매칭테이블을 주고받은 것은 더 강화된 암호화 방식인 SHA-512를 적용하려는 목적에서 제공받은 것으로 이후 삭제한 점 등을 들어 위와 같은 인식이나 의사를 인정하기 어렵다고 보았다.

마) 정리

이 사건은 비록 구법이 적용되던 시기에 발생한 사건이나 법원이 암호화조치의 적정성을 포함한 사실상의 가명처리의 당부에 관하여 논한 사례로서 주목받고 있음은 앞서 언급한 것과 같다.

민사법원이 암호화조치의 방식에 따라 적정한 비식별조치의 판단을 하려고 한 시도는 평가할 수 있다.[175] 다만, 위와 같은 각급 법원의 판단은 개인정보의 안전한 보호, 적정한 처리에 따른 이용의 측면에서 볼 때 아쉬운 면이 없지 않다. 물론 이와 같은 각 재판부의 판단은 구법 시행 당시 비식별조치, 식별가능성에 대한 법적 규율이 미흡하였던 점, 형사재판의 경우에는 고의의 인정여부가 핵심적 요소인 점 등에서 그 이유를 찾을 수 있다고 생각된다.[176]

174) 그러나 정보주체라고 볼 수 있는 환자, 처방 의사의 동의를 받은 사실이 없음은 명백하고, 약국 운영자들의 경우에도 진지한 동의를 하였다고 보기는 무리가 있다. 다만, 정보주체의 동의에 관한 논의는 본 연구의 범위를 벗어나는 것이어서 더 나아가 논하지는 않기로 한다.

175) 양기진, 앞의 논문, p.79 다만, 위 논문에서는 법원이 적절한 비식별조치가 이뤄진 경우 개인정보에 해당하지 않는 것처럼 여긴 것에도 의문을 표시(양기진, 앞의 논문, p.80)하고 있으나, 오히려 구법 규정대로라면 충분히 가능한 입론으로 보인다. 이와 같은 해석을 배제하기 위하여 신법에서는 비식별조치의 하나인 가명처리를 하였다고 하여도 일단 개인정보로 보겠다는 입장을 명확히 하였음은 몇 차례 살펴본 것과 같다.

176) 같은 맥락에서 구법 하에서는 제도의 경직성, 개인정보 여부 판단의 불확실성 등으로 인하여 '개인정보에 해당하는지 여부'를 직접 판단하기 어려웠으므로 위와 같이 손해나 고의의 유무로 해결할 수밖에 없었던 것으로 이해할 수 있다는 견해가 있다.(조성훈, 앞의 논문, p.245)

⑴ 부적절한 암호화조치

형사 재판부는 구체적인 판단을 명확히 하지 않았으나, 민사 제2심 재판부가 적절히 설시한 것과 같이 이 사건 암호화조치는 적정하다고 볼 수 없음이 명백하다.

1차 암호화의 경우 앞서 살펴본 것과 같이 단순한 치환에 불과하여 고전암호의 수준에 그친 것으로 암호문만으로도 빈도 공격(frequency attack), 전수조사 공격(exhaustive search attack)이 가능한데다가, 몇 개의 암호문-평문의 쌍이 주어지는 경우에는 일반인도 컴퓨터의 연산을 빌리지 아니하고 해독할 수 있을 것이다.[177] 따라서 현 시점에서 이와 같은 암호화조치는 적정한 가명처리 내지 비식별처리로 보아서는 아니 된다.

덧붙여 민사 제1심 재판부에서는 알파벳 치환 과정에서 암호문의 앞뒤로 1글자씩 추가하는 것을 노이즈라고 표현하였는데 이는 당사자의 주장을 만연히 받아들인 것으로 보인다.

즉, 가명정보 처리 가이드라인에 규정되었거나 수리암호학에서의 논하는 노이즈(noise)는 완전히 임의적(random)으로 선정, 부여하는 방식으로 암호화 알고리즘의 안전성을 높이는 것[178]인데, 사안에서의 노이즈는 일정한 규칙에 따라 이미 정해진 알파벳 중 하나를 고르는 것이어서 안전성에 기여하는 바는 거의 없다고 보아도 무방하므로 이를 노이즈라고 칭하는 것 자체가 암호화조치의 의미를 오해한 것이다.

2차 암호화의 경우에도 주민등록번호에 대한 해시함수를 이용한 암호화는 앞서 살펴본 것과 같이 그리 안전성이 높은 것이 아닌데다가, 결정적으로 개인정보처리자 상호 간에 평문(주민등록번호)과 암호문(해시값)의 쌍이 기재된 매칭테이블을 주고받은 사실 자체로, 더 이상 암호화조치의 성격을 상실하였다고 보아도 과언이 아니다.

---

177) 김명환, 앞의 책, p.31

178) 예를 들어 격자 기반 공개 키 암호에 쓰이는 LWE(Learning with error)문제의 경우 노이즈는 확률분포를 따르게 된다.(김명환, 앞의 책, p.211)

⑵ 구법 제18조 제2항 제4호의 적용 여부

민사 제2심 재판부는 위 조항은 적법하게 취득한 개인정보의 활용에만 적용될 수 있다는 해석을 하여 피고의 법 위반을 인정하였는데 명시적인 선례가 없는 가운데 적절한 결론이라고 생각하고, 이미 적정한 비식별조치가 이뤄지지 않았다고 보는 이상 위 조항을 적용할 여지도 없음이 명백하다.

덧붙여 구법은 '통계작성 및 학술연구 등의 목적을 위하여'라고 규정하고 있는데, 명백히 상업적 목적으로 개인정보를 매매한 본 사안에서 적용할 수 있는지는 심히 의문이다.[179] H회사와 미국 본사는 분명히 통계작성을 위하여 개인정보를 매입하였다. 그러나 그들은 취득한 정보를 시계열 통계로 작성한 후, 정부기관은 물론 기업에도 판매하였고, 이를 위하여 거액의 대가를 약학정보원에 지급하기도 하였기 때문이다.[180]

이러한 문제는 신법 시행 이후에도 말끔히 해결되었다고 보기는 어렵다. 법 제2조 제8호가 '과학적 연구'의 개념을 도입하여 이를 「기술의 개발과 실증, 기초연구, 응용연구 및 민간 투자 연구 등 과학적 방법을 적용하는 연구」로 정의하였고, 제28조의2 제1항에서 통계작성, 과학적 연구, 공익적 기록보존 등을 위하여 정보주체의 동의 없이 가명정보를 처리할 근거를 마련하였지만 이 사안과 같은 경우에 '과학적 연구', '통계작성'에 포함될 수 있을지는 이론[181]이 있을 수 있으므로 치열한 논의를 통하여 입법으로 해결함이 상당하다.[182]

---

179) 같은 취지로 GDPR의 규정 '정보주체가 합리적으로 기대할 수 있는 경우'에 본 사안이 해당한다고 보기 어렵다는 견해도 있다. (양기진, 앞의 논문, p.87)

180) 이와 같은 개인정보의 수수는 결국 약학정보원이 H회사에게 정보를 제공한 것으로 보아야 하고 처방내역이 민감정보인 이상 이는 동의나 법령상 근거가 있어야 한다는 일반론에서 이를 처리위탁으로 보았던 제1심 형사법원의 판단을 비판하는 견해도 있는데(선종수, 처방 정보 수집 및 제3자 제공에 관한 형사책임, 동아법학 제95호, p.106), 상당한 지적이라고 판단된다. 다만, 처리위탁이라고 하더라도 민감정보인 이상 비식별조치가 필요할 것이다.

181) 사견으로는 이 사안은 대가를 받고 처방정보를 제공한 것이어서 구법 당시의 과학적 연구에도 포함되지 않는다고 본다. (같은 취지의 견해, 선종수, 앞의 논문, p.109)

182) 이미 신법 개정 시 제안이유에서 '상업적 목적의 통계작성'이 삭제되었다고 한다. (김현숙, 앞의 논문, p.133)

(3) 문제점

이 사안은 모두 구법이 적용되던 시기에 발생하여 구법이 적용된 사안으로 신법은 물론 현재의 논의로 판결 내용 전부의 당부를 판단하는 것은 부적절할 수 있다.

그러나 아래와 같이 각 판결 내용과 본 연구에서 현재까지 논의한 내용을 비교할 때, 가명처리, 암호화조치에 관하여 어떻게 접근함이 상당한지, 신법 관련 어떠한 입법적 조치가 필요한지를 생각해 볼 수 있다.

(가) 적정한 가명처리, 암호화조치에 대한 인식 부족

우선 SHA-512라는 상대적으로 안전한 암호화 알고리즘이 사용되었다는 이유만으로 적정한 비식별조치가 있었다고 본 민사 1심 재판부의 경우, 식별가능성을 도외시한 견해로서 상당히 부적절하다. 매칭테이블이 존재하는 이상 그것이 수수되었는지, 개인정보처리자가 이를 보유하고 있었는지 여부를 불문하고 그 자체로 식별가능성이 강력히 인정될 수밖에 없다. 평문과 암호문의 쌍을 표시한 자료가 그대로 존재하는 이상, 쌍방향 암호화에서의 암호화키를 방치한 것이나 다름없기 때문이다.[183]

또한 형사 재판부들이 매칭테이블을 주고받은 것은 쌍방이 좀 더 안전한 암호화조치인 SHA-512를 적용하기 위함이었다고 보았으나, 이 역시 바로 위에서 논증한 것과 같이 암호화 알고리즘이나 비식별조치의 적정성을 이해하지 못한 것이다.

부연하면 입법자가 주민등록번호나 비밀번호에 대하여 일방향 암호화를 의

183) EU의 「적정한 임상 관행 및 임상 실험에 관한 지침(Directive 2001/20 of 4 April 2001 on the implementation of good clinical practice and the conduct of clinical trials)」에 따르면 임상실험 시 조사자는 임상결과를 제약회사나 각종 관련자에게 전달할 때 부호화된 형태만을 제공하고, 환자를 식별할 수 있는 일반 정보와 해당 부호에 결합되는 키는 따로 보관할 것을 규정하고 있다고 한다.(양기진, 앞의 논문, p.82) 아래에서 살펴보는 것과 같이 신법에서는 추가 정보에 관한 입법으로 일단 해결된 것으로 보이나, 매칭테이블의 수수와 같은 중요한 사정은 규범적 판단에 있어 정확히 평가되어야 할 것이다.

무로 부과한 것은 개인정보처리자도 이를 식별하지 못하게 하고자 하는 의도에 따른 것이다. 즉, 우리 법령은 위와 같은 정보가 암호문이 유출되더라도 사실상 복호화가 불가능하게 하여야 할 정보라고 보고 있는 것으로 이러한 정보에 대하여는 특별한 보호를 규정한 것이다. 그러므로 어떠한 암호화조치든 평문과 암호문 간의 매칭테이블을 작성하는 이상 실질적인 암호화조치의 강도는 급격히 낮아지므로 추가의 보안조치를 취하지 않는 이상 이는 부적절한 암호화조치에 해당한다는 결론에 이르러야 한다.[184]

그리고 비식별조치, 가명 조치, 암호화조치의 적정성은 여러 가지 사정을 고려한 식별가능성의 유무만으로 판단하는 것이 법의 문언은 물론 비식별조치 등의 취지에도 들어맞는 것인데, 일부 재판부에서는 재식별할 유인 내지 복호화할 동기가 없었다거나, 누구도 복호화를 시도하지 않았다거나 하는 등으로 행위자의 주관적 사정이나 사후에 추가로 일어난 사정으로 비식별조치 등의 적정성 여부까지 판단하려고 하고 있어 비식별조치 등의 개념, 중요성을 충분히 인식하지 못한 결과가 아닌가 한다.[185]

㈏ 적정한 비식별조치의 판단 기준의 필요성

특히 위에서 정리한 것과 같이 이미 어떠한 정보가 개인정보의 범위에 포섭될 것인가에 대하여는 다양한 판결이 있었고, 식별가능성의 문제는 곧 개인정보에 해당하는지의 문제에 연결된다고 하겠다.

즉, 누구의 기준에서 어떠한 기술, 비용을 들였을 때 어느 정도로 식별 가능한지가 식별가능성 인정의 기준이 될 것인데, 그러한 기준이 법령에 규정되지 아니한 결과 이 사안 관련 여러 판결 중 일부 내용에서 적정한 비식별조치에 관한 논증을 그르친 흔적이 엿보인다.

예를 들어 민사 1심 재판부에서는 제3자 입장에서는 암호문으로 개인을 식

184) 더구나 사안에서 쌍방이 매칭테이블을 주고받은 이상, 결국 어느 일방이 암호문으로 평문을 확인하려고 한 의도가 있었다고 넉넉히 추단할 수 있으므로 형사 법원과 같은 설시는 그 자체로 경험칙에 부합하지 않는다.

185) 양기진, 앞의 논문, p. 86에서는 이와 같은 법원의 판단에 식별가능성과 식별'감행'가능성을 혼동한 것으로 비판하고 있는데 적절한 견해라 하겠다.

별할 수 없는 점, 제3자가 암호문에 접근할 가능성이 없는 점 등을 들고 있고, 형사 재판부에서는 쌍방이 계약상 의무를 이행하는 과정에서 발생한 점, 매칭 테이블을 주고받은 것 자체만으로는 고의를 인정하기 어려운 점 등을 들고 있는데, 마치 판결들이 '제3자가 육안으로 보기에 식별할 수 없으면 식별가능성이 없는 것'이라고 기준을 제시하고 있다고 오해를 불러일으킬 소지가 크고, 실제 위와 같은 설시들은 암호화조치의 본질을 파악하지 못한 것에 이유가 있는 것으로 보인다.

즉, 제3자가 암호에 관하여 상당한 지식을 가지고 있고, 적절한 프로세서가 설치된 컴퓨터를 가지고 있으며 충분한 시간이 있는 경우에는 1차 암호화는 물론 2차 암호화도 복호화될 수 있는 점, 개인정보처리자 상호 간에 매칭테이블이 수수되어 각 개인정보처리자의 직원 등 내부자에 따라서는 충분히 복호화할 수 있는 점, 비식별조치는 개인정보처리자가 취한 조치, 예상되는 내외부의 공격자의 기술, 시간, 비용 등 객관적 사정에 따라 판단되어야 하는 점 등에서도 이 사건 비식별조치는 부적절하다고 볼 수 있을 것이다.

결국 이와 같은 규범적 판단의 난맥상이 발생하지 않도록 GDPR과 HIPAA에서 찾아볼 수 있을 정도로 가명처리, 암호화조치 등 비식별조치에 대하여 판단 기준을 제시할 필요가 있다.

⑷ 첨언 : 신법의 가정적 적용에 의한 해결

이상 살펴본 것과 같이 암호화조치에 의한 가명처리가 불충분하였다고 볼 수밖에 없으므로, 가명처리의 개념을 받아들인 신법에 의하더라도 정보주체의 동의가 없는 한 이와 같은 개인정보의 처리는 위법하고, 법 위반의 고의가 인정된다면 법 제17조, 제18조의 일반 규정 또는 가명처리에 관한 법 제28조의2 제1항에 따라 형사책임까지 추궁할 여지가 있다.

또한 신법은 제28조의2 제2항에서 가명정보를 제공할 때 특정 개인을 알아볼 수 있는 정보를 포함하는 것을 금지하고 있으므로, 사안에서 암호화 알고리즘을 공유하거나 매칭테이블을 수수한 것은 위 조항 위반행위로 볼 여지가 크다.

4. 개선 방안

신법은 가명처리, 가명정보의 개념을 도입하여 정보이용 주체의 개인정보 이용 가능성을 크게 확장하였고, 그 전제는 비식별조치인 가명처리, 암호화조치를 통한 개인정보의 보호에 있다고 할 것이다.

그러나 입법기술상의 문제점에 따른 입법의 흠결, 비식별조치에 대한 기술적 이해 부족으로 인하여 정보이용 주체가 비식별조치를 통한 적정한 개인정보 이용을 유도함은 물론 정보주체의 개인정보보호에 있어서도 여전히 장애가 존재한다고 판단된다.

이하에서는 법적 규율 및 법 해석상의 개선책이 어떠한지, 특히 가명처리, 암호화조치의 기술적 특성을 고려한 접근을 통하여 개인정보를 보호하고 이용할 수 있도록 하는 방안은 어떤 것이 있을지에 대하여 정리해 보고자 한다.

가. 입법론 : 규범력의 확보, 예측가능성의 보장

1) 가명처리 등의 판단 기준, 중요 개념의 입법 보완

본 연구를 통하여 신법이 가명처리, 가명정보의 개념을 도입하고도 법률은 물론 하위 법령 단위에서도 관련 개념, 요건, 특히 적정한 가명처리의 판단 기준을 구체적으로 정하지 않거나 그에 대한 규율을 회피한 결과, 가장 하위의 법적 규율 단위인 개보위의 고시에서조차 구법상의 표현인 '안전한 암호화 알고리즘', '상용 암호화 소프트웨어'와 같은 추상적 기준을 유지함에 그쳤고, 나머지 규율이 필요한 부분은 형식과 내용면에서 규범력을 부여할 수 없는 가이드라인에 의존하고 있음을 확인하였다.

이미 수차례 강조한 것과 같이 국가기관이 정보이용 주체에 대하여 적정한 가명처리 등이 이뤄지지 않았다고 일단 선언하는 이상, 정보이용 주체는 해당 개인정보를 이용하지 못하는 불이익을 받는 것에 그치지 않고, 민·형사상 책

임을 추궁당할 위기에 놓이게 된다. 따라서 수범자의 예측가능성을 보장하기 위해서라도 가명처리 등에 대하여는 여타의 행정법규와 마찬가지로 구체적인 행위 상황, 적법 요건을 규범력있는 법규명령이나 최소한 행정청에서 공표하는 행정규칙의 형식으로라도 규율할 필요가 있다.

그와 같은 시도에 대하여 기술중립성을 저해한다거나 기술 발전을 저해하는 불필요한 규제를 양산할 수 있다는 경계에 대하여는 동의하기 어렵다. 물론 가명처리 등의 내용의 상당 부분이 기술적인 것이고 법현실이 기술 발달에 따라 변동성이 있으므로 일정한 내용에 대하여는 가이드라인과 같은 유연한 방식을 취할 수밖에 없는 면에 대하여는 본 연구에서도 수긍하고 있다.

그러나 이미 살펴본 것과 같이 식별 주체(기준), 수단, 난이도 등과 같이 적정한 가명처리, 암호화조치의 판단 기준, 가명처리 등의 방식, 추가 정보의 의미와 요건 등은 일의적으로 규율할 수 있는 성격의 것들이고, 이미 GPPR과 HIPAA와 같은 국외의 입법례에서 성공적으로 규율, 적용하고 있기 때문이다.

이와 같은 기준을 확립함으로써 행정청은 일관된 규제 행정을 시행할 수 있고, 수범자는 자신의 기술 개발, 이용에 대하여 진정한 의미의 가이드라인을 법령에서 구할 수 있으며, 특히 수사기관, 사법기관의 자의적 규범적 판단을 사전에 예방하는 효과가 기대된다.

구법이 적용되던 시기에 내려진 각급 법원의 판결의 내용을 검토하여 보면, 식별주체에 대해 일관된 판단이 없었고, 사건의 결론에 맞추어 식별가능성에 대한 판단을 그르치거나 회피하여 수범자들에 대하여는 예측가능성을 보장할 수 없었고, 정보주체인 일반 국민들에게는 자신의 개인정보가 제대로 보호되지 못한다는 인상을 남기게 되었다.

가) 적정한 가명처리, 암호화조치의 판단 기준의 제시

우선 개인정보보호법에서 적법한 가명처리의 요건을 개괄적으로 언급하고, 이를 하위 법령에 명시적으로 위임하면서 아래와 같은 내용을 입법으로 반영할 필요가 있다.

첫째 식별가능성의 판단 주체를 명문으로 규정하되, 개인정보처리자에 더하

여 합리적 범위 내의 제3자까지 확장하고, 제3자의 범위에 대하여는 정보처리 과정, 정보의 내용, 추가 정보 등 관련 정보의 소재 등을 고려함을 규정할 필요가 있다.

둘째, 식별 수단의 수준에 대하여도, 정보처리시점 또는 최신의 시점으로 사용한 가능한 기술 및 시간, 비용 등을 종합할 때 합리적으로 예상 가능한 수단을 식별 수단으로 명확히 표현할 필요가 있다.

셋째, 이와 같은 요소에 더하여 개인정보 처리의 성격, 범위, 내용, 목적, 정보주체에 대한 영향 및 앞서 HIPAA가 전문가의 관점[186]을 제시할 때 들고 있었던 요건인 '통계학적, 과학적 기법의 사용' 및 개인정보 자체의 정보와 합리적으로 사용가능한 정보 등을 식별가능성 판단에 있어 추가적으로 고려할 요소로 정할 수 있을 것이다.

이와 같은 정도의 입법적 조치는 식별가능성의 존부라는 규범적 판단에 관한 추상적 요건에 지나지 않아, 기술중립성을 해할 정도로 기술적인 내용을 담고 있지 않고, 법관 등의 자유심증을 해치거나 구체적 타당성을 염두에 둔 사법, 행정작용을 제한할 정도로 구체적인 기준이라고 보기는 어려우므로 충분히 실현가능한 입법이다.

법률의 단계에서 식별가능성의 판단 주체 및 식별 수단에 대하여 개략적으로 정하고 나머지 요소들은 대통령령 이하 법령에 위임함으로써 현실을 적시에 반영할 수 있도록 할 수 있을 것이다.

나) 중요한 개념, 요건의 상세한 규율

⑴ 추가 정보

가명처리에 사용되는 추가 정보는 매우 중요한 개념으로 각종 법률효과가 발생하는 법률요건으로 심지어 형사책임의 요건에까지 이어진다. 예를 들어 추가 정보를 별도로 보관, 관리하는 등의 조치를 취하지 아니하여 개인정보를

186) 전문가의 평가 내용만으로 식별가능성을 정하는 것은 법관의 자유심증에 따른 재판제도를 헌법상의 결단으로 취하고 있는 우리 법제에서 그대로 이를 받아들이기는 어려울 것이다.

분실・도난・유출・위조・변조 또는 훼손당한 경우에는 법 제73조 제1호에 따라 2년 이하의 징역 또는 2천만 원 이하의 벌금에 처해지는데, 추가 정보의 개념에 대하여는 법률, 시행령은 물론 안전성 확보기준에 관한 고시를 다루고 있지 아니하므로 앞서 인용한 것과 같이 명확성의 원칙에 반한다는 지적이 있을 정도로 큰 문제점이 있다고 하겠다.

개보위의 가이드라인에서 추가 정보를 「개인정보의 전부 또는 일부를 대체하는 가명처리 과정에서 생성 또는 사용된 정보로서 특정 개인을 알아보기 위하여 사용·결합될 수 있는 정보(알고리즘, 매핑테이블 정보, 가명처리에 사용된 개인정보 등)」라고 설명한 것이 일응의 입법안으로 볼 수 있겠다. 이와 같은 정도의 입법 내용으로는 명확성의 원칙을 충족하면서도 기술중립성을 해하지 않는다고 판단된다.

⑵ 익명정보

익명정보의 경우 가명정보에 비하여 상대적으로 상세한 요건을 법률단계에 두었으나, 법에 정해진 「시간・비용・기술 등을 합리적으로 고려할 때 다른 정보를 사용하여도 더 이상 개인을 알아볼 수 없는 정보」의 요건 중, '다른 정보', '개인을 알아볼 수 없는'의 부분은 중요한 개념임에도 그 의미를 추측할 단서가 없으므로, 대통령령 이하의 법령에 위임하고 예시 등을 들어 설명이 불가피하다.

이 때 다른 정보나 식별가능성은 위와 같이 적정한 가명처리의 판단 요건이 정해지게 된다면 상당 부분의 판단 요건을 차용하되 식별불가능을 요구하는 방식을 취할 수 있을 것이다.

⑶ 과학적 연구

신법에서 과학적 연구의 개념을 정의한 것은 큰 의미가 있다고 할 것이나, 법 문언상의 「기술의 개발과 실증, 기초연구, 응용연구 및 민간 투자 연구 등 과학적 방법을 적용하는 연구」라는 표현만으로는 약학정보원 사건에서와 같은

처방행태의 분석이나 환자와의 관련성을 연구하면서도 상업적 목적으로 통계처리를 하게 되는 경우도 포함되는지 여부와 같이 다양한 사례에서 명쾌한 답을 주지는 못한다.

'과학적 연구'는 통계작성, 공익적 기록보존과 함께 법 제28조의2 제1항, 법 제28조의3 제1항에서 가명처리가 가능한 경우 및 가명정보의 결합이 가능한 경우에 대한 법률요건이므로 과학적 연구에 포섭되지 아니함에도 가명처리하거나 가명정보를 결합하게 되면 법 제71조 제2호 또는 제4의3호에 따라 형사처벌까지 가능한 상황이다.

결국 정의 규정에서 언급된 요소가 인정되는 경우에도, 상업적 목적의 처리도 가능한 것인지, 순수한 연구실 내의 연구만 허용하겠다는 것인지 등을 알 수 없으므로 최소한 하위법령의 단계에서 산업 분야별, 이용 형태별 예시를 들고 범주화하는 노력이 필요하다.

⑷ 법 제15조 제3항 등에 따른 이용의 세부 규율

덧붙여 신법이 도입한 또 다른 형태의 정보주체의 동의 없는 개인 정보이용인 법 제15조 제3항, 제17조 제4항의 경우에도 비록 시행령이 GDPR의 규정을 참조하여 나름대로 상세한 규율을 하였으나, 당해 규율의 기술적 부분에 대한 내용은 고시 등에 위임하여 규범력을 확보하고, 개보위의 가이드라인에서 최소한 가명처리, 암호화조치에 관한 부분에 관하여는 위 조항 부분도 포함하여 설명하고 있음을 명백히 하면 좋을 것이다.

⑸ 가명처리와 암호화조치 간의 관계 설정

현재 법령에서 가명처리와 암호화조치의 관계를 규정한 부분은 확인되지 아니하고. 암호화조치는 이미 구법에서부터 제한적으로 규정되어 온 점, 시행령 제14조의2 제1항 제4호에서 법 제15조 제3항 등에 따른 이용의 요건을 규정하면서 유일하게 「가명처리 또는 암호화」라는 표현을 사용하고 있는 점 등이 확인된다.

그러나 본 연구에서 검토한 암호화조치와 가명처리의 개념에 비추어 볼 때, 암호화조치는 가명처리의 가장 유용한 방안으로 봄이 상당하고, 실제 개보위의 가이드라인에서는 가명처리의 여러 형태 중 '개인정보의 삭제 또는 대체' 중 하나의 방안으로 암호화조치를 들고 있다.

한편, 구법에서부터 암호화조치의무가 규정되었던 고유식별정보, 주민등록번호, 비밀번호에 대하여는 법개정 이후에도 여전히 암호화조치가 의무로서 부과되고 있으므로, 이들 정보에 대하여는 암호화조치 이외의 다른 가명처리기법은 사용될 수 없을 것이다.

따라서 위와 같이 가명처리의 요건을 좀 더 상세히 규정하는 부분에 한하여 암호화조치가 가명처리의 여러 기법 중 하나임을 밝혀두는 것으로 양자의 관계를 명백히 할 수 있다고 본다.

2) 법 규정 상호 간, 의료법 등 단행법과의 관계 정리

가) 민감정보에 대한 가명처리 여부

이미 논증한 것과 같이 일단 현행법의 해석상으로도 민감정보를 가명처리의 대상에서 제외할 이유는 없다고 본다. 다만, 민감정보가 구법에서부터 특별한 보호를 받아왔고 앞으로도 그와 같은 보호는 유지하여야 할 것으로 보이므로, 이 부분에 대하여는 법 제23조의 민감정보 관련 부분이나, 법 제28조의2 가명정보의 처리 관련 부분에 민감정보 또한 가명처리를 통하여 이용할 수 있다는 취지의 확인적 성격의 조항을 신설하는 것이 정보이용 주체에게 법적 안정성을 부여하는 것이라 하겠다.

나) 의료법 등 단행법과의 관계

이미 법 제6조에서 신법은 개인정보에 관한 규율에서 다른 법률의 규정이 우선함을 선언하였으므로 의료법에서 개인정보에 관한 규율이 있다면 이에 의하여야 함은 본 연구에서 논증하였다.

그러나 법은 개인정보의 안전한 보호와 적정한 이용을 염두에 두고 제정, 개정된 일반법으로 그러한 입법 목적에 따라 규율하고 있으나, 개별 단행법은 각자의 입법 목적, 규율 대상에 집중하여 입법된 것이어서 개인정보의 보호, 처리를 통한 활용에 충분하고도 적절한 해법을 제시하지 못하는 경우가 예상된다.

또한 의료정보에 관하여 논증한 것과 같이, 모든 의료정보가 의료법의 규율을 받는다고 단정할 수 없다. 의료법과 같은 개별 단행법은 수범자나 규율 대상을 특정하고 있으므로 그러한 법률이 모든 국민을 규율하고자 한다면 해당 법조에서 이를 명기하여야 한다.

요컨대, 우선 입법자는 개인정보의 처리를 통한 활용이 필요한 분야의 개별 단행법을 일괄하여 검토하고 개인정보보호법에서 규율할 여지가 있는지에 따라 개별 단행법과의 충돌 논란을 입법 단계에서 정리할 필요가 있는데, 대표적인 예가 본 연구에서 언급한 의료기관이 아닌 곳에 보유, 관리하여 공공데이터 형식으로 제공되는 의료정보일 것이다. 나아가 의료법과 같이 개별 단행법이 선점하여 명시적으로 개인정보의 이용을 제한하고 있는 분야에 대하여도 전문가 집단의 연구, 국민의 공감대 형성을 통하여 개인정보보호법의 규율에 따라 적정한 개인정보의 이용을 구현할 수 있도록 전향적인 조치를 검토할 때가 되었다고 생각한다.

3) 시행령, 고시의 보완

이미 수차례 지적한 것과 같이 법이 주요한 내용이 추가되는 형식으로 개정되었음에도 시행령은 대부분의 사항을 구법 당시의 내용을 그대로 유지하고 있을 뿐인 고시에 재위임한 상태이다.

이상 입법 개선이 이뤄지는 경우 시행령, 고시에 상당한 내용이 추가될 것으로는 보이는데, 특히 가명처리, 가명정보, 익명정보, 추가 정보, 과학적 연구 등 신법이 새롭게 도입한 개념에 대하여는 반드시 별도의 법률 요건을 두어 규율하는 것이 상당하다고 본다.

4) 가이드라인에 대한 전면적 검토

가) 가이드라인의 한계, 필요성

본 연구에서 가이드라인의 법적 성격, 지위, 효과에 대하여 상세히 살펴본 이유는 법현실상 분쟁이 발생할 소지가 높은 분야임에도 가이드라인이 법령에서 제대로 다루지 않은 부분을 상세히 설명하고 있기 때문이다.

그 결과 법령의 위임은 물론 근거조차 없는 가이드라인이 행정지도와 유사한 형식을 차용하여 수범자에게 특정한 조치를 권장, 설명이라는 외관을 취한 채 사실상 의무로 부과하고 있는 이상 이는 행정행위의 측면에서나 법적 규율의 측면에서 온당하지 않다는 결론을 도출하였다.

물론 개인정보보호 분야, 특히 가명처리, 암호화조치 분야는 상당한 논의 내용이 기술에 관한 것이고, 기술의 발달은 물론 기술의 적용 분야가 급격히 변화하고 있어, 일정한 단계에서부터는 최소한 자율규제를 대체[187]하는 일종의 공적 규율로서 가이드라인의 필요성은 인정한다.

나) 상위 법령에 위배되지 않는 가이드라인

그러나 본 연구에서는 가이드라인이 신법에서 도입한 가명처리 등을 해석함에 있어 입법취지는 물론 문언상 의미에도 반하는 내용을 제시하고 있는 예를 다수 확인하였으므로 이러한 부분은 만큼은 충실한 검토 후에 우선적으로 시정하여야 한다.

예를 들어 「가명처리 수행 당시의 목적, 처리 환경은 적정한 가명처리를 판단하는 일응의 기준이 될 수 있을 뿐임에도 이러한 요소로 가명정보의 이용이 제한된다.」는 것(가명정보 처리 가이드라인)은 법률에 근거가 없을 뿐만 아니라 신법의 개정취지를 몰각하는 설명이다. 앞서 정리한 것과 같이 가명처리 수행 당시의 목적, 처리 환경에 가명처리를 구속하여 얻는 효과를 알 수가 없

187) 이원복, 앞의 논문, p.212

고, 이와 같은 문언은 법률에 의하여 보장된 재산권의 행사인 가명처리를 부당하게 제한하는 것이다.

또한, 「식별자를 암호화조치 하여도 이를 삭제하여야 한다.」는 설명(보건의료데이터 처리 가이드라인)은 암호화조치의 성격을 오해한 것으로서 특히 법령에 따라 가명처리한 가명정보를 가이드라인이 불충분한 근거를 들어 삭제를 권장하고 있는 것이어서 법률 규정에 반하는 것이다.

특히 적정한 가명처리에 대한 구체적인 설명도 없이 「안전한 가명처리가 개발되지 않은 경우 가명처리 자체를 금하는 것(보건의료데이터 처리 가이드라인)」 또한 법률의 근거 없이 정보이용 주체에게 사실상의 의무를 부과하고 있는 것이다.

그리고 입법으로 해결하여야 할 문제, 민감정보의 가명처리, 의료법과의 관계 등을 가이드라인이 설명하는 것은 그 자체가 무용한 것이고 오히려 현장의 혼선을 야기할 뿐이므로 입법으로 해결함과 동시에 가이드라인은 상위 법령의 취지에 맞추어 정리되어야 할 것이다.

다) 가이드라인이 필요한 부분에 대한 가이드라인 제시

오히려 본 연구에서는 가이드라인이 진정한 의미의 가이드라인으로서 수범자들에게 가이드라인을 제시할 수 있는 부분을 일부 확인하였는데, 추가 정보의 예시, 적정한 관리 방법, 암호화 알고리즘별 장단점 등이 그것이다. 이는 아래에서 항을 바꾸어 보완 가능한 예시를 간략히 지적해 보겠다.

각 가이드라인이 가명처리 또는 그 대상에 관하여 설명하는 정도의 수준으로 위와 같은 부분에 대하여 일응의 의견을 제시해 준다면 수범자들에게 의미있는 자료를 제공할 수 있다고 본다.

라) 법령상 근거의 마련

앞서 강조한 것과 같이 행정기관이 가이드라인을 설정할 수 있도록 입법상 조치를 통하여 법률 또는 하위 법령에서 구체적 사항을 정하여 가이드라인에

위임하거나, 최소한 제정의 근거를 제시하는 것은 이미 다수의 예가 있다.

만일 법령에서 구체적 사항을 정하여 위임하는 것이 가이드라인에도 법규성을 부여하게 되어 기술중립성을 저해한다거나 탄력적인 규제 행정이 어려워지게 될 것이 우려된다면 여성가족부장관의 양육비 가이드라인과 같이 법률상으로는 근거만을 마련하는 것도 하나의 방법임을 제안한다.

나. 입법취지에 따른 법령의 해석, 적용

자유심증주의 등 법관의 직업적 양심에 따른 재판을 보장하는 법치국가의 헌법 체제 하에서 고의, 과실의 판단 기준까지 개인정보보호법과 같은 단행법이 제시하여서는 아니 된다. 그러나 개인정보보호법의 해석, 적용에 있어 개인정보처리자에 대한 고의, 과실의 인정 여부는 법 위반행위의 여부, 인식에 초점을 맞추어야 한다고 본다.

그런데 앞서 살펴본 일부 판결이 고의나 과실을 인정하지 않는 결론을 설명하는 과정에서, '결과적으로는' 개인정보가 유출된 사실이 없다거나, 행위자들이 '계약상 의무를 이행'하려고 하였다거나, 위반행위 이후에 매칭테이블을 폐기하였다거나, '최초 수집 목적'을 벗어나지 않았다거나, '경제적 이익을 기도한 것으로는 보이지 않는다'는 등 설시에서 나타나는 사정은 행위 이후에 발생한 사실 또는 정황사실에 불과한 것으로 이러한 사정을 종합하여 행위의 위법성을 부인하는 것은 개인정보보호법의 입법 취지를 도외시하는 법의 해석, 적용이라고 생각한다.

즉, 사기죄의 편취의사나 계약상 채무불이행에 대한 과실과 같이 여러 가지 정황 사실을 밝혀내고 이를 종합하여 심리적 구성요건 요소를 인정하는 것과는 달리 개인정보보호법에 대한 위반행위는 법에서 요구하는 의무가 발생한 상황, 그럼에도 이를 이행하지 않는다는 인식(예를 들어 개인정보를 처리하는 상황을 알고도 가명처리를 하지 않는다는 인식, 안전성 확보조치를 취하지 않는다는 인식) 또는 금지된 행위를 한다는 인식(예를 들어 추가 정보를 넘겨준다는 인식) 등으로 고의를 인정하고, 민사상 과실의 경우에는 그러한 의무의

불이행, 위반행위 경위 자체에 초점을 맞추어야 할 것이다.

그 이유는 개인정보보호법이 개인정보자기결정권이라는 정보주체의 헌법상 권리를 보호하기 위한 입법목적도 엄연히 존재하고 있으므로 법 위반행위는 적극적으로 발견하여 개인정보 보호에 노력할 필요가 있는 점, 위 법은 일응의 행정법규로서 개인정보보호에 관한 각종 의무를 부과한 후, 그 위반으로 인한 개인정보의 침해가 발생할 위험을 방지하는 것에 목적이 있는 점, 법 적용 여부는 문언대로 해석, 적용하되 행위자의 동기, 실제 개인정보 침해의 결과는 형사의 경우 양형사유, 민사의 경우 손해 인정 유무, 범위의 산정에 충분히 반영할 수 있는 점 등을 들 수 있을 것이다.

덧붙여 개인정보보호의 측면에서 정보이용 주체의 사정을 폭넓게 고려하여 민형사상 책임을 면하게 하거나 감경하여 주는 사법기관의 판단이 늘어나게 되면, 사회 구성원 대다수를 차지하는 정보주체들의 입장에서는 이를 국가기관의 온정적 대응으로 여기고 각자의 개인정보 이용에 부정적인 태도를 가지게 되어 결국에는 개인정보 이용의 활성화에 부작용을 가져올 우려를 배제할 수 없음을 밝혀 두고자 한다.

다. 기술적 측면 : 암호화조치의 적극적 활용

구법에서부터 특정한 개인정보에 관하여는 개인정보 이용주체에 대하여 암호화조치의무가 부과되었고, 신법이 도입한 가명처리의 구현에 있어 암호화조치는 상당한 비중을 차지하고 있다고 할 것이다.

위에서 살펴본 것과 같이 암호화조치를 구현하는 암호화 알고리즘은 역사적으로는 수학이론 및 컴퓨터 관련 기술의 발전에 따라 개발되어 왔고, 이용 환경에 따라 다양한 형태가 존재하고 있으므로, 개인정보의 보호, 이용의 측면을 함께 고려하여 각각의 장단점을 따져보아야 한다.

개인정보보호법을 개정하게 된 가장 주된 이유는 정보이용주체가 정보주체의 동의 없이 적법, 적정하게 개인정보를 처리하여 이를 이용하도록 하겠다는 것에 있는데, 4차 산업혁명 시대의 개인정보 처리는 대부분 디지털데이터의

형태로 변환된 개인정보가 대상이고, 정부나 기업이 관심을 가지고 기술의 연구, 개발에 몰두하는 분야는 그와 같은 디지털데이터가 대량으로 빈번하게 처리되는 소위 '빅데이터' 관련 사업이다.

이와 같은 현실을 고려할 때, 수리암호를 응용한 암호화조치는, ① 수학적 난제에 기한 암호화 알고리즘의 경우 권한 없는 자가 복호화를 시도할 경우 수학적 난제를 푸는 만큼의 어려움이 보장되어 안전성이 있고, ② 컴퓨터 연산 기능 향상에 따라 효율적으로 암호화가 가능하며, ③ 복호화 또한 알고리즘으로 정해져 있어 권한이 부여된 정보처리자가 정보주체를 재식별하여 처리결과를 제공할 수 있고, ④ 특히 위와 같이 개인정보 처리 방식이 디지털데이터를 이용한 컴퓨터 연산인데, 수리암호를 이용한 암호화조치도 알고리즘의 형태로 구현되어 소프트웨어로서 사용가능하므로 빅데이터에 대한 가명처리가 용이하며, ⑤ AES, 해시함수, 동형암호 등 유용한 암호화 알고리즘마다 고유의 특성, 장단점이 있어 데이터 형태, 처리환경 및 목적에 따라 다양한 선택이 가능하다는 장점이 있다.

다만 수리암호에 따른 암호화 알고리즘을 사용하더라도 비밀키의 존재에 따른 복호화 위험성, 평문과 암호문을 대응시키는 매칭테이블, 암호화 알고리즘의 유출과 같이 정보처리에 관하여 적법한 권한을 가지고 있는 사람과 그로부터 권한을 위임받은 사람의 권한 남용, 컴퓨터 기술을 활용한 해킹과 같이 내외부로부터의 공격에 대한 취약요소가 상존하고 있는 것도 사실이다.

따라서 이와 같은 소위 재식별의 위험을 사전에 예방할 수 있는 입법과 그에 따른 적정한 법의 해석, 적용이 필요하다. 이와 같이 암호화조치 등 기술적 측면에서 가명처리에 대한 접근 방식을 간략히 살펴보겠다.

1) 해시함수 등 일방향 암호화에 대한 적정한 규율

구법에서부터 주민등록번호, 비밀번호 등에 대하여는 '일방향 암호화'를 의무로 부과하고 있었는데 이는 위와 같은 정보가 정보주체에 대하여 중요한 의미를 지니고 있고 개인정보보호에 직결된다고 보았기 때문이다.

현재 일방향 암호화 알고리즘의 대표적인 예는 위에서 상세히 살펴본 해시

함수로서 널리 쓰이는 만큼 다양하고도 빈번한 공격에 노출되어 있기도 하다. 특히 해시함수의 중요한 특징인 일방향성, 충돌회피성을 신뢰할 수 있어야만 적정한 가명처리로서의 암호화 알고리즘이라고 볼 수 있을 것인데 그에 관한 기술적 규율이 미진하다.

먼저 사전 계산을 통하여 예상 가능한 암호문을 마련하고 공격을 시도하는 외부자에 대하여는 사전 계산량을 늘이기 위해서라도 솔트를 가미하여 해시함수를 사용하여야 하고, 실제 업계에서 널리 사용되고 있기는 하다.

그러나 개보위 가이드라인에서는 솔트의 개념만 설명하고 있을 뿐 구체적인 내용, 요건을 언급하지 않고 있다. 이러한 경우 불성실한 정보이용주체가 솔트를 사용하지 않은 채 개발 이후 상당한 시일이 경과한 MD4, SHA1 등의 함수를 사용하고도 최소한 법령이나 가이드라인 상으로는 특별한 문제가 없다고 주장할 여지가 있다.

따라서 정보처리자 등에 대하여 해시값을 사전에 계산하여 공격을 시도하는 경우를 상정하고 평문에 추가의 수치를 더하여 해시함수를 적용하여야 함을 고시의 단계에서 언급하고, 상세한 기술적 내용은 가이드라인에서 충분히 다뤄야 할 것이다.

마찬가지로 일방향암호화를 요구하는 가장 중요한 이유인 복호화를 통한 재식별가능성을 원천적으로 봉쇄할 필요가 있는데, 예를 들어 약학정보원사건에서와 같이 평문과 암호문을 대응한 '매칭테이블'을 제작한 경우, 정보처리자나 그 위임을 받은 사람이 정보주체를 쉽게 재식별할 수 있고, 극히 일부의 매칭테이블이 유출되기만 하면 공격자는 해시함수의 형태, 솔트의 존재, 평문의 형식을 충분히 추론해 낼 수 있기 때문이다.

물론 이와 같은 행위가 「원래의 상태로 복원하기 위한 추가 정보를 별도로 분리하여 보관, 관리하는 등의 조치를 취하지 아니하여 개인정보를 분실·도난·유출·위조·변조 또는 훼손당한 경우」에 해당된다면 신법 제73조 제1호, 제28조의4 제1항에 따라 형사 처벌될 수는 있을 것이다.

그러나 개인정보의 보호에 만전을 기하고자 한다면 최소한 일방향암호화가 요구되는 경우에 대하여는 매칭테이블과 같은 추가 정보의 생성, 보관 자체를 원칙적으로 금지함이 상당하다. 이미 일방향암호화를 하는 이상, 이를 평문으

로 복호화할 이유가 없는 것이고, 복호화를 할 필요가 있는 정보라면 좀 더 엄격한 요건을 부과한 후에 AES 등 양방향암호화 알고리즘을 사용하도록 하는 것이 법의 취지는 물론 개인정보보호의 측면에도 부합한다고 생각한다.

2) 암호화 알고리즘 등 가명처리 기법에 대한 가이드라인 등 법적 설명

개보위의 가이드라인에서 살펴본 것과 같이 개보위는 다양한 가명처리 기법을 소개하고 있으나, 안전성이 전혀 보장되지 않는 경우, 부적절하거나 불충분한 설명에 그친 경우 등이 다수 확인되고 있다. 이러한 부분들은 가명처리, 암호화조치에 관한 적정한 이해를 통해 상세히 정리할 필요가 있다.

가) 삭제, 부분 삭제, 마스킹과 식별가능성의 문제

개인정보를 완전히 삭제하는 것은 개인정보보호의 측면에서는 양호한 조치로 보이나, 개인정보의 처리 결과를 정보주체에 돌려주는 것이 불가능하고, 개인정보를 삭제하면 개인정보의 이용 자체가 어려워지는 결과를 초래할 수 있고, 이에 따라 정보이용주체가 스스로의 판단에 따라 이를 선택하게 될 것이다. 문제는 부분 삭제, 마스킹 등으로 개인정보를 일부 삭제하게 되는 경우에는 식별가능성에 유의할 수밖에 없는데, 식별가능성의 판단 기준이 명확하지 않은 현행 법제상에서는 개인정보보호의 측면에서 우려를 낳게 된다.

즉, 식별자를 원천적으로 삭제하는 외의 방법은 그 자체로 식별가능성의 우려가 해소되지 않을 여지가 있고, 정보 이용에도 제약이 심하다고 할 것이다.

나) 정보이용주체가 선택할 가능성이 희박한 가명처리 기법

데이터의 수치가 특이한 경우, 이를 통하여 개인정보가 재식별될 우려가 있다는 이유로 수치 자체를 올림, 버림, 반올림하거나 평균값으로 대체하는 기법이 라운딩, 범주화 등으로 볼 수 있는데, 이러한 기법은 결국 수치에 주목하게

되는 빅데이터 처리에서는 쉽게 활용할 수 없는 것이므로 정보이용주체가 이를 택하지 아니하면 무용지물이 될 수밖에 없다. 순열(치환)의 경우에도 일정한 알고리즘에 따라 순열, 치환을 하게 되는 경우 공격자의 공격에 급격히 취약하게 된다. 반대로 데이터의 재배열을 무작위로 하게 된다면 데이터와의 정보주체 간, 데이터 상호 간의 연관성이 훼손되므로 이용가치도 저감될 것이다.

이러한 가명처리의 기법도 일정한 규칙, 즉 알고리즘에 따라 이뤄지게 되면 수리암호를 이용한 암호화 조치에 비하여 안전성이 떨어지게 되고, 위와 같이 정보이용 주체가 정보를 이용할 실익도 낮아져 가명처리 기법으로서 유용한지 의문이 제기될 수밖에 없다.

다) 잡음(노이즈) 추가 요건

이미 약학정보원 사건에서 드러난 것과 같이 일정한 정수를 노이즈에 추가하는 것은 개인정보의 보호에는 별다른 효과가 없음에도, 개보위의 가이드라인에서는 이와 같은 사례를 예시로 열거하고 있을 뿐이다.

앞서 지적한대로 수리암호학의 개념에서 언급되는 노이즈는 확률분포를 따르는 등으로 안전성을 높이고 있고, 그러한 노이즈를 사용하는 이상 더 이상 단순한 노이즈 추가가 아니라, 격자 암호, 동형암호 등과 같이 최신의 수리암호화 알고리즘의 세계로 접어들게 되는 것이다.

그러한 수준의 안전한 노이즈 추가가 아닌 한, 이를 가이드라인에서 설명하고 언급하게 되면 불성실한 개인정보처리자에게 부적정한 가명처리를 적법하다고 주장할 수 있는 빌미를 줄 것이라는 식의 우려는 앞서 유사한 논점에서 수차례 지적하였다.

라) 정리 : 암호화 알고리즘별 장단점 및 처리 정보 대상에 따른 유형화 필요

이상 종합하여 검토할 때, 법령과 가이드라인에서 암호화조치 의무를 상세히 규율하지 않은 것은, 암호화조치와 가명처리의 관계가 자명함에도 이를 명확히 처리하지 않은 점, 암호화조치가 구법에서부터 규정된 탓에 신법이 가명

처리를 도입하게 되었음에도 일단 기존 고시 내용을 유지하고 있는 점, 개인정보의 보호 및 이용에 암호화조치가 안전성과 효율성을 동시에 추구할 수 있는 수단이 암호화조치라는 것을 주목하지 않은 점, 같은 맥락에서 정보처리의 대상에 따른 가명처리 등의 특성을 깊게 고찰하지 않은 점 등에서 이유를 찾을 수 있다.

이러한 문제의식에서 암호화조치의 추가적인 규율 방안을 생각해 보면, 앞서 검토하였던 보건의료데이터 처리 가이드라인에서 긍정적인 예를 찾을 수 있다고 생각한다. 위 가이드라인은 상정 가능한 다양한 보건의료데이터의 형태, 식별가능성, 처리 방식에 따라 가명처리 기법의 장단점 내지 적합한 가명처리의 예시를 들고 있기 때문이다.

즉, 가명처리 기법 중 암호화조치가 가지는 장점 중 하나는 데이터의 형식, 처리 환경 및 방식의 측면에서 가장 적합한 암호화 알고리즘을 선택할 수 있다는 것에 있으므로 위와 같이 암호화조치에 대하여 상세한 규율을 하게 되는 경우에는 처리 대상 정보의 특성에 따라 다양한 암호화조치기법을 설명하고, 각자의 장단점을 제시하는 방식이 효과적이라고 생각한다.

예를 들어 정보주체에게 정보처리 결과를 돌려주게 되는 서비스의 경우, 그 처리과정에서 정보주체의 개인정보임을 식별할 수 없는 암호화조치인 동형암호와 같은 기법이 가장 최적화 된 것이고, 특별한 처리는 필요하지 않으나 정보주체를 식별한 상태로 정보를 보관할 필요가 있다면 AES와 같은 쌍방향 암호화가 사용되어야 할 것이며, 처리 대상 정보의 재식별가능성을 가장 경계하는 경우라면 해시함수와 같은 일방향 암호화가 필요할 것이고, 정보주체 간의 차이를 식별할 수 있으면서도 정보주체에 따른 통계학적 처리가 필요하다면 차분 프라이버시와 같은 대안을 생각해 볼 수 있겠다.

3) 개인정보의 유출 등 사고 발생 이후의 대비 필요

GDPR에서 주목할 만한 내용은 일단 가명처리, 암호화조치가 된 정보 그 자체가 유출되는 등의 사고가 발생하였을 때에도 식별가능성에 관심을 두고 있다는 것으로, 전문 제85조가 앞서 정리한 것과 같이 불충분한 가명처리 상태

에서 당해 정보가 유출된 경우에는 재식별의 위험성이 높으므로 정보이용주체의 관리를 벗어나게 이후에도 안전성을 확보할 의무를 별도로 규정하고 있기 때문이다.

이와 같은 태도는 주목할 필요가 있는 것이, 암호화 등 비식별조치가 이루어진 정보라도 일단 유출되게 되면 공격자가 이를 보유한 상태에서 시간이나 메모리의 제한 없이 식별을 시도하게[될 것이기 때문이다.

그러나 현행 개인정보보호법은 GDPR과는 달리 개인정보의 유출 등 사고가 발생한 경우, 개인정보처리자에 대하여 정보주체에 대한 통지의무를 부과할 것인지, 그 요건은 어떠한지 등에 대한 규율에 그치고 있는 맹점을 찾아볼 수 있다.

따라서 비식별조치의 적정성을 판단할 때, 특히 기술적인 측면에서 유출된 이후에도 합리적인 시간, 방법, 비용을 들이더라도 재식별이 어려운지 여부를 고려하여야 할 것이고, 이와 같은 문제를 경계하는 입법 내지 행정적 조치가 필요하다.

4) 암호화조치에 대한 규범적 판단 기준

가) 식별가능성 판단 기준의 강화

이미 본 연구에서 정리한 것과 같이 개인정보의 이용을 보장하기 위해서라도 그 전단계인 비식별조치에 대하여는 좀 더 엄격한 잣대를 요구할 수밖에 없다고 본다.

구법 시대의 일부 판결례에서 나타나듯 사법기관에서는 사안에 가장 상당한 결론을 내리기 위하여 비식별조치의 내용, 적정성 외에도 다양한 정황사실을 고려하여 종합적 판단을 내린 것으로 보인다. 판결의 과정에서는 구체적 타당성을 외면할 수 없다는 현실적 고충을 이해하나, 그와 같은 논증 과정이 반복되면 수범자들은 비식별조치, 즉 암호화조치 등 기술적 조치가 하나의 고려요소에 불과하다는 오해를 가질 우려가 크다.

즉, 위와 같은 암호화조치에 대한 규범적 판단의 태도가 지속되는 경우, 정

보이용 주체는 굳이 수리과학적으로 안전성이 입증된 암호화조치를 취할 필요성이 없다고 여기면서, 다소 부실한 암호화 알고리즘을 사용하더라도 자신에게 피해에 대한 예측가능성이 없었다거나 법위반의 고의가 없었다는 등으로 면책 주장을 하고자 하는 유혹을 뿌리치기 어려울 것이다.

요컨대 사법, 행정기관이 적정한 비식별조치, 가명처리 등을 판단함에 있어 기술적인 측면을 가장 중요한 요소로 상정하여야 한다. 세부적으로는 암호화 조치의 내용, 특성, 구현 방식, 잠재적 공격 가능성 및 그 수단 등을 면밀히 고찰하여 당해 암호화 조치의 안전성, 신뢰성이 현재 기술 및 합리적인 시간, 비용을 고려하였을 때에도 인정할 수 있을지 여부로 기본적인 판단을 내릴 수 있어야 할 것이다.

그에 따라 적정한 암호화조치가 이루어졌다면 그 외에 법령에서 임의적 조치로 규정한 조치가 이행되지 않았다고 하더라도 비식별조치 등의 적정성을 인정할 여지를 넓힐 수 있을 것이고, 반면 적정한 암호화조치가 이루어지지 않았다는 판단을 하는 경우, 그 외에 취해진 다른 조치에 대하여는 그 효용성에 대하여 보다 엄격하게 판단할 수 있어야 할 것이다.

나) 기술적 요소를 고려한 규범적 판단

본 연구에서 제안한 각종 입법 조치가 이뤄지게 되면 규범적 판단의 법적 요건이 좀 더 명료해질 것으로 기대되므로, 사법, 행정기관이 적정한 비식별조치 등에 대한 규범적 판단을 내리게 되는 경우에는 기술적 요소를 폭넓게 고려할 필요가 있다.

가이드라인의 법적 지위에서 검토한 것과 같이, 기술중립성의 요청에 비추어 보아도 기술적 요소야말로 성문 법령으로 규율하기에는 적절하지 않은 것이 사실이므로 규범적 판단의 단계에서는 반드시 사안별로 기술적 요소, 즉 암호화조치에 있어서는 개별 암호화 알고리즘의 구현 방식, 장점 및 취약점을 충분히 검토하여야 할 것이고, 그 과정에서는 학계, 실무계의 전문가의 의견을 필요적으로 청취함이 상당하다는 결론에 이르렀다.

이와 같은 절차 내지 고려를 보장하기 위해서라도 적정한 비식별조치의 판

단 기준을 법제화할 때 기술적 요소를 고려하여야 한다는 정도의 주의적 규정을 두게 된다면 규범적 판단을 하게 되는 주체가 기술적 요소에 좀 더 비중을 두고 개별 사안에 접근하게 될 것으로 예상한다.

Ⅴ. 결론

1. 가명처리, 암호화조치의 중요성

본 연구의 서두에서 언급한 것과 같이 정보주체인 사회 구성원들이 각자의 개인정보가 스스로의 동의 없이 제3자가 활용하는 것에 대하여 공감대를 가지려면, 개인정보가 적정히 처리되어 정보주체의 개인정보자기결정권이 침해될 우려가 없다는 확신을 주어야 할 것이다.

법 개정논의가 절정을 이루던 시기의 여론조사를 인용한 언론보도에 따르면 한 여론조사의 경우 응답자의 80%가량이 법 개정 추진 사실을 알지 못하였고, 동의 없는 개인정보의 수집·이용에 반대하는 의사를 표시하였으나, 또 다른 여론조사의 경우 응답자의 70%가 기술개발·신산업육성 등에 데이터 활용이 필요하고, 84%가 공익 목적에 기여하는 경우 개인정보를 제공할 의향이 있다고 각각 답하였다고 한다.[188]

이러한 일견 상반되는 것처럼 보이는 여론조사 결과는 조사자의 편향, 조사기법의 차이에서 비롯된 것이라고 여길 수도 있으나, 정보주체들이 개인정보의 활용에 관하여 가지고 있는 인식의 양면을 모두 드러내는 것으로 충분히 양립가능한 결과라고 본다. 즉, 정보주체의 동의 없는 개인정보의 처리 등 활용은 입법 조치에 더하여 사회적 공감대 형성이 필요하다고 볼 수밖에 없다.

본 연구에서 살펴본 것과 같이 신법의 개정에 따라 정보주체의 동의 없이 정보이용 주체가 개인정보를 처리하고자 할 때에는 가명처리가 사실상 필수적

188) 민주노총 등이 2019. 10. 1,000명을 대상으로 실시한 여론조사에서는 개인정보보호법 개정 추진을 모른다는 답변비율이 81.9%, 동의 없이 가명정보를 수집, 이용하는 것에 반대한다는 답변비율이 80.3%로 산출되었다는 기사(연합뉴스, “데이터3법 깜깜이 논의 ‘성인 5명 중 4명 꼴 개정 추진 몰라’”, 2019. 11. 3.자.)가 있고, 한편 경기연구원이 2020. 2. 1,000명을 대상으로 실시한 여론조사에서는 위와 같이 기술개발 등에 데이터 활용이 필요하다는 답변비율이 70%, 공익 목적에 기여할 경우 (개인정보를) 제공할 의향이 있다는 답변비율이 84%로 산출되었다는 기사(뉴스1, “국민 84% ‘데이터 3법, 공익 목적 기여 시 정보제공 의향 있다.’” 2020. 3. 4.자)도 있다. 이상 김현숙, 앞의 논문, p.135의 각주 52, 53에서 재인용

요건이 되었다고 할 수 있다. 가명처리를 도입한 이유는 효율적인 개인정보의 처리, 활용에만 있는 것이 아니라 개인정보 보호의 강화엔 있기 때문이다.

또한, 가명처리에는 다양한 기법이 있을 수 있으나, 개인정보에 직접 연결되는 식별자를 삭제하는 방법 외에는 정보 처리의 보안성, 효율성 및 활용 가능성의 측면에서 수리암호를 이용한 암호화조치가 유용하다는 결론에 이르렀다.

2. 정보이용 주체에 대한 예측 가능성 부여

개인정보보호법이 개정된 이유는 일정한 요건 하에 적정한 가명처리 등의 방안을 통하여 정보주체의 동의 여부로부터 자유로운 정보이용을 보장하고자 하는 것에 있음은 수차례 살펴본 것과 같다.

본 연구를 통하여 가명처리, 가명정보, 추가정보, 익명정보 등 주요한 법적 요건에 대하여 법률은 물론 하위 법령에서 구체적인 설명이 이뤄지지 아니한 점, 그러한 설명을 담당하고 있는 가이드라인은 법적 근거가 희박함에도 상위 법령의 규정과 배치되는 설명 또는 불충분한 설명에 그치고 있는 점 등을 확인할 수 있었다.

본 연구의 문제의식은 일단 국가기관이 정보이용 주체에 대하여 적정한 가명처리 등이 이뤄지지 않았다고 선언하게 되면, 정보이용 주체가 해당 개인정보를 이용하지 못하는 불이익을 받는 것에 그치지 않고, 민·형사상 책임을 추궁당할 위기에 놓이게 되므로, 정보이용 주체와 같은 수범자의 예측가능성을 보장하기 위해서라도 가명처리 등에 대하여는 여타의 행정법규와 마찬가지로 구체적인 행위 상황, 적법 요건을 규범력있는 법규명령이나 최소한 행정청에서 공표하는 행정규칙의 형식으로라도 규율할 필요가 있다는 것이다.

이와 같은 조치에 포함시킬만한 내용의 예시, 대안을 위와 같이 정리해 본 결과, 정보이용주체에 대하여 반드시 필요한 규율의 예를 다수 확인할 수 있었고, 이를 구체적이고도 규범력있는 수단으로 강제할 때, 오히려 정보이용주체에 대하여 책임감 있고도 적극적인 개인정보 이용을 주문할 수 있을 것이라는 결론에 이르렀다.

3. 정보주체의 정보이용에 대한 신뢰 확보

정보이용 주체가 정보주체로부터 별도의 동의를 받지 아니하더라도 일정한 요건에 따라 개인정보를 처리, 이용할 수 있도록 길을 열어준 것이 개정 개인정보보호법의 주된 개정 취지라고 강조하게 되면 정보주체의 자기 정보의 이용, 관리에 대한 의사는 뒷전에 밀리는 것이 아니냐는 비판이 제기되는 것은 자연스럽다.

그러나 정보주체의 자기정보결정권이 헌법상 권리로 확인된 가운데, 정보주체가 자기정보결정권을 소위 구체적 권리로서 소구할 수 있는 다양한 방안이 이미 여러 가지 형태로 개별 법령에 규정되어 있다. 따라서 정보주체가 개인정보의 처리, 이용의 안전성, 보안성 등에 의구심을 가지게 되는 이상, 정보주체는 민형사절차 등을 통하여 개인정보의 처리 등에 문제를 제기할 수 있을 것이다.

요컨대 개인정보보호법이 개정되었다는 것만으로 개인정보의 부당한 침해가 흔해질 것이라거나, 개인정보의 자유로운 이용이 완전히 보장되었다고 볼 수는 없는 셈이다. 에 의구심을 가지는 이상 각종 법적 분쟁이 반복될 수밖에 없으므로 법률의 개정만으로 개인정보의 자유로운 이용이 보장되기는 어렵다고 본다.

결국 이와 같은 사정을 고려하면, 신법이 정보주체의 동의 없는 개인정보의 이용을 제도화한 이상, 개인정보보호에 관한 법령의 해석, 적용은 좀 더 엄격해 질 필요가 있고, 본 연구에서 기존 판결례의 해석론이나 입법의 흠결 내지 불비를 지적한 이유가 여기에 있다.

개인정보의 이용에 대한 부당한 제한이 아닌 이상, 개인정보보호에 관한 각종 법제에 대하여는 늘 보완할 사항을 검토함이 상당하고, 그러한 법제를 개별 사안에 적용할 때에는 개인정보의 침해 위험성을 염두에 두고 엄격히 판단하는 태도가 필요함을 강조하고 싶다.

4. 맺음말

입법으로 새로운 현상을 규율하는 것은 입법자가 해당 현상을 제대로 반영하지 못하였다는 비판, 현상에서 제기되고 있는 문제점에 대한 적절한 해법을 제시하지 못하고 있다는 비판이 늘 제기되므로 어려운 일이다.

2020. 4. 데이터 3법이 개정된 것은 개인정보가 포함된 디지털데이터의 처리에 있어 구법 등 기존의 법제가 이를 제대로 뒷받침하기는커녕 관련 기술의 개발, 실현으로 나아가는 여정의 여러 요소에서 걸림돌이 되었다는 반성에서 비롯된 것이므로, 이에 따라 개정된 신법은 개인정보의 안전한 보호를 주된 목표로 여전히 가지고 있으면서, 한편으로는 개인정보의 적정한 처리를 통한 활용을 보장, 장려하여야 한다는 새로운 과제를 안고 있다.

그러나 본 연구에서 살펴본 것과 같이 신법 체제 하의 법령, 가이드라인에도 개인정보보호의 측면은 물론 개인정보이용의 측면에까지도 여전히 보완할 점이 없지 않고, 특히 법의 해석, 적용의 측면에서 기술적 특성을 충분히 고려하지 않은 일부 판결례의 문제점이 반복될 염려가 있어 앞으로 학계, 산업계 및 관련 정부기관이 함께 노력하여야 할 것이다.

본 연구 전반에서 반복하여 강조한 것은 개인정보의 안전한 보호와 적정한 처리를 통한 활용이 양자택일의 문제가 아니라 함께 추구할 수밖에 없는 목표라는 관점이다. 개인정보가 안전하게 관리되고 있다는 사회적 공감대가 형성될 때에 개인정보의 적정한 처리를 통한 활용이 보장될 수 있을 것이다.

개인정보의 보호와 이용은 결국 헌법상 기본권의 보장, 법률상 권리의 행사의 문제로 귀결되므로 이와 관련된 입법상의 조치가 충실히 이루어졌을 때, 정보이용 주체 등 수범자에 대하여 예측가능성을 부여하고, 법을 해석, 적용, 집행하는 행정, 사법기관이 일관되고도 적법한 조치를 취할 수 있을 것이다.

본 연구에서 지적한 논증한 내용들이, 입법자에 의하여 충실한 입법 보완으로 이어지고, 행정청, 사법기관들이 가명처리와 암호화조치에 대한 깊은 이해를 토대로 부적절한 개인정보 관리를 시정하면서 적정한 개인정보의 이용에 대하여는 적극 보호, 장려하게 되는 작은 계기가 될 수 있기를 바란다.

<참고문헌>

1. 단행본

김명환, 수리암호학 개론, 경문사, 2019
박노형, 개인정보보호법, 박영사, 2020
고학수 등 7인, 인공지능 시대의 개인정보 보호법, 박영사, 2022
개인정보보호위원회, 가명정보 처리 가이드라인, 2022
보건복지부 등, 보건의료데이터 활용 가이드라인, 2022
국무조정실 등 4개 부처, 개인정보 비식별조치 가이드라인, 2016

2. 논문, 기사

계인국, 이성엽, 보건의료 데이터 활용의 법적 쟁점과 과제, 공법연구 제50집 제2호, 2021. 12.
김송옥, 가명정보의 안전한 처리와 합리적 이용을 위한 균형점 -데이터 3법에 대한 헌법적 평가를 겸하여- , 공법연구 제49집 제2호, 2020. 12.
김현숙, 과학적 연구목적을 위한 개인정보 처리에 관한 비교법적 연구, 정보법학 제24권 제1호, 2020. 1.
김희정, 가명정보 미동의 처리의 기본권 침해 검토, 법학논총 제45권 제1호, 2021. 1.
박노형·정명현, 빅데이터 분석기술 활성화를 위한 개인정보보호법의 개선방안 - GDPR과의 비교 분석을 중심으로, 고려법학 제85권, 2017. 1.
선종수, 처방 정보 수집 및 제3자 제공에 관한 형사책임, 동아법학 제95호, 2022. 5.
양기진, 개인정보의 범위에 관한 연구 - GDPR의 비식별조치와 약학정보원 사건의 검토 ……, 선진상사법률연구 통권 제84호, 2018. 1.
오길영, 데이터 비식별화 정책에 대한 규범적 비판, 공법연구 제46집 제2호, 2017. 12.
이상엽, 망 이용계약 가이드라인 법적 함의와 전망, KISO저널 제37호, 2019. 12.

이석배, '보건의료데이터 활용 가이드라인'의 현행법상 문제점, 대한의료법학회「의료법학」제22권 제4호, 2022. 1.
이원복, 유전체 연구와 개정 개인정보 보호법의 가명처리 제도, 이화여자대학교 법학논집 제25권 제1호, 2020. 1.
정영진, 보건의료데이터와 개인정보 보호와의 관계에 대한 소고, 법학논총 제34권 제3호, 2022. 2.
조성훈, 개인정보보호와 형사책임 : 가명정보 특례와 목적의 합리적 관련성을 중심으로, 법학평론 제12권, 2022. 4.
천지영, 노건태, 데이터 3법 시대의 익명화된 데이터 활용에 대한 제언, 정보보호학회논문지 제30권 제3호, 2020. 6.
뉴스1, "국민 84% '데이터 3법, 공익 목적 기여 시 정보제공 의향 있다.'" (2020. 3. 4.)
연합뉴스, "데이터3법 깜깜이 논의 '성인 5명 중 4명 꼴 개정 추진 몰라'", (2019. 11. 3.)
Apple Inc., Apple Platform Security Guide (May, 2022.)
Alan Marcus, "Data and the fourth industrial revolution", World Economic Forum (https://weforum.org/agenda/2015/12/data-and-the-fourth-industrial-revolution, 2021. 11. 8. 확인)
Bellare, Mihir 등 3명, Message Authentication using Hash Functions— The HMAC Construction, p.1 Appears in RSA Laboratories' CryptoBytes, Vol. 2, No. 1, (Spring 1996)
Cheon, Jung Hee 등 4명, Homomorphic encryption for arithmetic of approximate numbers(2017)
Daniel Solove, "Why I love the GPPR", http://teachprivacy.com/why-i-love-the-gdpr (2022. 7. 17. 확인)
Fenton, James L. 등 3명, NIST Special Publication 800-63-3, p.54. NIST Technical Series Publications (June 2013)
Jan H. Samoriski 등 3명, Encryption and the first amendment (Mar. 23, 2009)
Jeri Clausing, U.S. Losing Battle on Control of Data Encryption, Study

says, The New York Times, (Feb. 9, 1998)
Joel Brinkley, U.S. Eases Encryption Software Export Bans, The New York Times (Sep. 17, 1998)
Jung Hee Cheon and 6 others, Privacy Preserving COVID-19 Contact Tracing with Homomorphic Encryption. International Conference on Appropriate Technology(ICAT), 2021
Kristin Lauter; Michael Naehrig; Vinod Vaikuntanatha, Can Homomorphic Encryption be Practical?, p.114, CCSW '11: Proceedings of the 3rd ACM workshop on Cloud computing security workshop (October 2011)
Laurence D. Smith, Cryptography, The science of Secret Writing(1942)
Google Removing SHA-1 Support in Chrome 56, Threat Post (Nov. 16. 2011.)

3. 웹페이지

https://foreignaffairs.com/articles/2015-12-12/fourth-industrial-revolution (2021. 11. 8. 확인)
https://forbes.com/sites/bernardmarr/2016/04/05/why-everyone-must-get-ready-for-4th-industrial-revolution(2021. 11. 8. 확인)
https://www.law.go.kr/법령/개인정보보호법, 법률 제16930호 제정·개정 이유(2022. 10. 24. 확인)
https://www.law.go.kr/법령/신용정보의이용및보호에관한법률/ 법률 제16957호 제정·개정 이유(2022. 10. 24. 확인)
https://www.law.go.kr/법령/정보통신망이용촉진및정보보호등에관한법률/ 법률 제16955호 제정·개정 이유(2022. 10. 24. 확인)
https://eprint.iacr.org/2006/294(2023. 1. 3. 확인)
https://www.merriam-webster.com/dictionary/encryption(2021. 11. 1. 확인)
cryptograms-Sha256 : Fast, pure and practical SHA-256 implementation, hackage.haskell.org(2022. 5. 15. 확인)
https://www.techopedia.com/definition/5066/international-mobile-equipme

nt-identity-imei(2023. 1. 1. 확인)
https://www.igi-global.com/dictionary/usim-universal-subscriber-identity-module/31282(2022. 11. 1. 확인)

4. 국내외 판결례 등

헌법재판소 2005. 5. 26.자 99헌마513호 결정
헌법재판소 2021. 11. 25 자 2017헌마1384, 2018헌마90, 145, 391(병합) 결정
대법원 1995. 2. 14. 선고 94누12982호 판결
대법원 2009. 12. 24. 선고 2009두7967호 판결
서울중앙지방법원 2011. 2. 23. 선고 2010고단5343호 판결
서울중앙지방법원 2015. 1. 14. 선고 2014고단5061호 판결(제1심), 같은 법원 2015. 4. 9. 선고 2015노387호 판결(제2심)
서울중앙지방법원 2016. 7. 15. 선고 2015고합336호 판결(제1심), 서울고등법원 2020. 1. 31. 선고 2016노2150호 판결(제2심)
서울중앙지방법원 2017. 9. 11. 선고 2014가합508066호 등 판결(제1심), 서울고등법원 2019. 5. 3. 선고 2017나2074963호 등 판결(제2심)
서울중앙지방법원 2020. 2. 14. 선고 2015고합665호 등 판결(제1심), 서울고등법원 2021. 12. 23. 선고 2020노628호 판결(제2심)
수원지방법원 성남지원 2017. 9. 15. 선고 2017고단1438호 판결(제1심), 수원지방법원 2018. 4. 12. 선고 2017노7275호 판결(제2심)
대전지방법원 논산지원 2013. 8. 9. 선고 2013고단17호 판결
Reno v. ACLU, 512 U.S. 844(1997)
Sorrel v. IMS Health Inc. 564 U.S. 552(2011)
IMS Health Inc. v. Mills, 616 F.3d 7(2009)
IMS Health Inc. v. Ayotte, 550 F.3d 42(2010)

Abstract

– A research on the Pseudonymisation in the Personal Information Protection Act of the Republic of Korea –

Sung, Kibum
Master of Science in Digital Forensics
Department of Mathematical Information Science
The Graduate School of Convergence Science and Technology
Seoul National University

On Feb 4th, 2020, there were several revisions in the Personal Information Protection Act ('PIPA') of the Republic of Korea.

The revision is said to be aiming to boost and enlarge the legitimate process and use of personal information without the consent of the subjects of that information('the subject'), which will lead to research and developments of new technologies related to big data such as artificial intelligence, and machine learning.

On the other hand, since the Constitutional Court of the Republic of Korea affirmed that the subject of personal information retains the constitutional right to decide whether to disclose, use or access one's personal information, the PIPA also can be one of the legal measures to secure such constitutional right.

Thus the primary purpose of this research is to inquire into the practical balance between the sound protection of personal information and the legitimate and effective use of personal information.

The new PIPA introduced the notion of pseudonymization which is set up for pseudonymized information so that the processor of the personal information("the processor") came to be allowed to use and process the personal information without the consent of the subject.

Encryption via mathematical cryptography is one of the most efficient, secure

pseudonymization methods. Since it can be performed easily through software in computing systems, mathematical cryptography has specific strength in the big data processing. In addition, it is a secure way of encryption in that it is based on mathematical conundrums. The processors can also choose from a variety of algorithms, such as AES, homomorphic encryption, and hash functions, depending on shapes, processing environments, and purposes of the data.

There are several flaws in legal implements under the new PIPA, however: (1) We don't see some important concepts, and explanations such as the criteria of appropriate pseudonymization, or additional information in pseudonymization from the law itself and its subordinate statutes; (2) As a result, such important concepts and standards are only being explained by guidelines of several related government agencies which are not formal statutes, so that the processors cannot figure out easily whether their process of personal information is legitimate or not; (3) Such guidelines mislead the processors and the subjects with irrelevant explanations not based on the statutes or contradict to them; (4) Furthermore, we still have several legal issues even with the new PIPA, such as the relationship between the Medical Act and the PIPA, and whether the 'sensitive information' in the PIPA also can be pseudonymized under the law.

There are foreign legislative examples for the revised PIPA, one is the European Union's General Data Protection Regulation ('GDPR') and the other one is the US federal's Health Insurance Portability and Accountability Act of 1996 ('HIPAA').

The GDPR clearly regulates the criteria for determining identifiability: (1) It enlarges the viewpoint of judging identifiability to a third party with a reasonable possibility in addition to the processor; (2) To determine whether the information is identifiable or not, it directs to regard the state of the art, cost and time used for the identification with the character, scope, content of the processing.

The HIPAA also gives some perspectives on the concept of 'identifier'. It clarifies the definition of an identifier and demands the processor erase the whole identifier. It allows the processor to use de-identification as an alternative if the experts with statistical and scientific techniques give the opinion that the

probability of re-identification is very low.

There were discussions on de-identification measures, even in the old PIPA era, since there were attempts by the processors and some rulings that tried to exclude some personal information from the scope of the law, which had undergone de-identification such as pseudonymization.

However, such trials sometimes led to improper conclusions. For example, one court used the fact that two processors exchanged a matching table that paired plaintext and cipher, to deny their intention of malpractice. Other courts misunderstood a mere substitution as a legitimate de-identification. Those rulings were derived from the insufficient stipulation of the statutes concerning important concepts including de-identification, and pseudonymization.

One reason for the drawbacks we can find in the judgments above lies in an attempt to allow the use of personal information through non-identification measures by an insufficient guideline since the old PIPA didn't regulate the concept and requirements of de-identification. In sum, the main problem was that the standards for the judgment of the court had not been prepared by legislation.

Inconsistent and sometimes inappropriate decisions and judgments by government agencies and courts will not provide legal stability to the processors and are likely to cause anxiety about the use of personal information by data subjects.

This study concluded that the new PIPA, subordination statutes, and related guidelines still need improvements to pseudonymization and encryption measures: (1) criteria for determining proper pseudonymization or encryption measures; (2) the concept and requirements of additional information used for pseudonymization; (3) whether or not to pseudonymize sensitive information; (4) the relationship with other legislation such as the medical law, etc.; and (5) examples of pseudonymization or encryption techniques.

Especially, for guidelines, it is necessary to establish a legal basis by a delegation from the higher statutes, correct conflicts with them, and specify technical matters that guidelines can explain, such as the advantages and disadvantages of encryption algorithms and proper management of additional

information.

Processing of personal information using big data presupposes frequent calculations on large amounts of digital data, thus encryption measures based on mathematical cryptography that can be implemented in software are very useful.

However, since statutes and guidelines do not clearly address the technical aspects of encryption measures, it is necessary to categorize the advantages and disadvantages of each encryption algorithm regarding the form or field of applicable information.

Accordingly, government agencies and courts should scrutinize the characteristics of encryption algorithms and concentrate on judging the legality of encryption algorithms used as de-identification measures.

It is important to secure the safety and legitimacy of the processing of personal information for mitigating the concerns of the subjects of the information so that members of society can support the professor's research and development of new technology on big data.

Expecting follow-up research and legislation will continue on the problems pointed out in this research, and the proposed contents, through these attempts, government agencies will achieve harmony between technology development and personal information protection with an appropriate interpretation and application of the PIPA.

Keywords: PIPA, de-identification, identification, pseudonymization, encryption, mathematical cryptography, personal information
Student Numbers: 2021-28895