서울대학교
인공지능정책
이니셔티브

개정 개인정보 보호법상
가명정보의 개념 및 가명처리에
관하여

이슈페이퍼 2020-1

데이터 3법 시대의 과제:

- 가명처리
- 연구목적 활용
- 차등적 프라이버시
- 데이터 거래

http://sapi.co.kr/

개정 개인정보 보호법상 가명정보의 개념 및 가명처리에 관하여

고학수 서울대학교 법학전문대학원 교수
백대열 서울대학교 법과대학원 박사과정
구본효 서울대학교 법과대학원 박사과정
정종구 서울대학교 법과대학원 박사과정
김은수 서울대학교 아시아태평양법연구소

요약문

‘데이터 3법’의 개정 내용 중 핵심적인 것 하나는 가명정보의 개념을 명시적으로 도입한 것이다. 가명정보는 넓게는 개인정보로 분류되지만, 별도 유형의 특수한 개인정보로 규정되어 정보주체의 동의 없이 통계 목적, 과학적 연구 목적, 공익적 기록보존 목적 등을 달성하기 위하여 이용하거나 제공할 수 있다. 또한 개정법은 가명정보의 결합에 대한 규정을 도입하고 있는데, 결합은 전문기관을 통해 수행해야 하는 것으로 규정되었다. 가명정보와 관련된 법제도의 변화는 이론적으로나 실무적으로나 적지 않은 새로운 과제를 제시한다. 여러 관련 쟁점들에 대한 상세하고 진지한 논의를 통해 데이터 시대로의 변화에 적극적으로 대응할 필요가 있다.

1. 서론

데이터 경제로의 패러다임 변화를 반영하여 관련 법령을 정비하는 것이 시급하다는 반복된 논의를 기반으로, 2020년 1월 9일에 개인정보 보호법[01], 정보통신망 이용촉진 및 정보보호 등에 관한 법률, 신용정보의 이용 및 보호에 관한 법률[02] 개정안(이하 ‘데이터 3법’)이 국회 본회의를 통과했다. 새로이 개정된 데이터 3법은 2월 4일에 공포되어 오는 8월 5일부터 시행될 예정이다. 데이터 3법과 관련해서는 많은 논의가 있어왔지만, 직접적으로는 2018년에 대통령 직속 4차 산업혁명위원회 주관으로 관계부처, 시민단체, 산업계, 법조계 등의 전문가가 참여하여 두 차례 열린 ‘해커톤’에서의 논의결과에서 촉발된 것이다.[03]

정부의 설명으로는 개인정보 보호법의 주요 개정내용은 다음과 같다.[04] ① 가명정보 개념을 도입하고, ② 동의없이 처리할 수 있는 개인정보 활용방안을 마련하였으며, ③ 개인정보의 범위를 명확히 하고, ④ 개인정보처리자의 책임성을 강화하며, ⑤ 개인정보 보호체계를 일원화하였다는 점이다. 정보통신망 이용촉진 및 정보보호 등에 관한 법률의 주요 개정사항은, 이 법에 규정되어 있던 개인정보 보호에 관한 사항들을 삭제하고 이를 개인정보 보호법에 이관하는 것이다.[05] 신용정보의 이용 및 보호에 관한 법률(이하 ‘신용정보법’)의 주요 개정내용은 다음과 같다고 한다.[06] ① 가명정보 개념을 도입하고, ② 신용평가회사의 라이센스를 세분화하며, ③ 신용정보업의 업무범위를 확대하고, ④ 마이데이터 산업을 도입하며, ⑤ 소비자의 정보인권 보호관련 규정을 추가한 것이다. 총체적으로는, 데이터 3법 개정의 핵심은 가명정보와 관련된 규정들을 명시적으로 도입한 것 그리고 집행권한 및 조직의 상당부분을 개인정보보호위원회로 일원화한 것에 있다고 볼 수 있다. 그 이외에 목적합치의 원칙을 도입한 것도 중요한 변화이다.[07]

---

01 개인정보 보호법은 2018. 11. 15. 인재근 의원이 의원입법으로 개정안을 대표발의한 이후 1년여 기간이 지난 2019. 11. 27. 국회 행정안전위원회를 통과했다. 2020. 1. 9. 국회 법사위의 전체회의를 통과했으며 2020. 1. 9. 국회 본회의를 통과하였다. 2020. 2. 4. 개정법률안이 공포되었으며 6개월 후인 2020. 8. 5. 시행될 예정이다.

02 신용정보의 이용 및 보호에 관한 법률은 2018. 11. 15. 김병욱 의원이 의원입법으로 개정안을 대표발의한 이후 1년여 기간이 지난 2019. 11. 28. 국회 정무위의 법안1소위를 통과하고 2019. 11. 29. 국회 정무위의 전체회의를 통과했다. 2020. 1. 9. 국회 법사위의 전체회의를 통과한 후 같은 날 국회 본회의를 통과하였다. 2020. 2. 4. 개정법률안이 공포되었으며 6개월 후인 2020. 8. 5. 시행될 예정이다.

03 행정안전부, 데이터 규제 혁신, 청사진이 나왔다, 보도자료, (2018. 11. 22.).

04 관계부처 합동 브리핑, 개인정보 보호법 개정 후속조치 계획, (2020. 1. 21.).

05 정보통신망 이용촉진 및 정보보호 등에 관한 법률 개정이유 및 주요내용 (법률 제16955호. 2020. 2. 4. 일부개정).

06 금융위원회, 신용정보법 설명회, (2020. 2. 20.).

07 이동진, “목적합치의 원칙과 가명정보의 특례”, 법률신문 연구논단, (2020. 3. 23.).

이 글에서는 가명정보의 개념 및 가명처리에 관한 규정들이 데이터 3법 특히 개인정보 보호법에 어떤 형태로 포함되었는지 살펴보고, 향후 어떤 이슈들에 대해 추가적인 논의가 필요할지 분석해 보기로 한다.

2. 개정 데이터 3법상 가명정보와 가명처리의 개념

가. 개인정보 및 가명정보의 개념

개인정보의 개념과 관련하여, 기존의 개인정보 보호법상으로는 식별의 가능성 및 결합의 용이성이 핵심적인 요건이었다.[08] 즉, '개인을 알아볼 수 있는 정보'가 개인정보의 법적 개념이 되고, 만일 다른 정보와 결합하여 개인을 알아보게 되는 경우에는 다른 정보와 '쉽게 결합하여' 알아보게 되는 경우에만 개인정보인 것으로 규정되었다. 개정법에서 규정하는 개인정보는 기존 법의 요건 중에서 결합의 용이성을 판단함에 있어 입수가능성 등을 추가로 고려하도록 정하였다.[09] 즉, 다른 정보와 쉽게 결합하여 개인을 알아볼 수 있는지 여부를 판단함에 있어, 해당 정보의 '입수 가능성 등 개인을 알아보는 데 소요되는 시간, 비용, 기술 등을 합리적으로 고려하여야 한다'고 규정하였다.

개정법은 가명정보를 개인정보의 특수한 한 형태인 것으로 규정하였다. 즉, 가명처리를 통하여 '원래의 상태로 복원하기 위한 추가 정보의 사용·결합 없이는 특정 개인을 알아볼 수 없는' 상태가 된 정보를 가명정보라고 정하고, 이를 개인정보의 일부로 나열하였다.[10] 이때 가명처리란, '개인정보의 일부를 삭제하거나 일부 또는 전부를 대체하는 등의 방법으로 추가 정보가 없이는 특정 개인을 알아볼 수 없도록 처리하는 것'을 말한다(개정 개인정보 보호법 제1의2호).[11] 한편 개인신용정보의 맥락에서, 가명처리란 추가 정보를 사용하지 아니하고는 특정 개인인 신용정보주체를 알아볼 수 없도록 개인신용정보를 처리하는 것이고, 가명정보란 가명처리한 개인신용정보를 말한다(개정 신용정보의 이용 및 보호에 관한 법률 제2조 제15호 및 제16호).

개정된 개인정보 보호법은 개인에 관한 정보를 실질적으로 3가지 유형으로 분류하였다. 첫째, 식별정보 내지 식별가능한 정보이다. 이는 성명, 주민등록번호 및 영상 등을 통하여 개인을 직접적으로 알아볼 수 있는 정보와, 해당 정보만으로는 특정 개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 알아볼 수 있는 정보를 합하여 규정하는 것이다.[12] 두 번째 유형의 정보는 가명정보이다. 이는, 위의 2가지 유형의 정보를 가명처리 함으로써 원래 상태로 복원하기 위한 추가 정보의 사용·결합 없이는 특정 개인을 알아볼 수 없게 된 정보를 말한다. 셋째, 명시적으로 규정한 것은 아니지만 익명정보의 개념이 있다. '시간·비용·기술 등을 합리적으로 고려할 때 다른 정보를 사용하여도 더 이상 개인을 알아볼 수 없는 정보에는 개인정보 보호법을 적용하지 아니한다'라고 규정한 것을 통해 익명정보를 간접적으로 규정한 것으로 해석할 수 있는 것이다.[13] 개정 신용정보법의 경우에는, '더 이상 특정 개인인 신용정보주체를 알아볼 수 없도록 개인신용정보를 처리하는 것'을 익명처리라 규정함으로써, 익명정보의 개념을 실질적으로 도입한 것으로 해석될 수 있다(개정 신용정보법 제2조 제17호).

나. 가명정보의 활용과 취급

개정 개인정보 보호법에는 개인정보의 활용과 관련하여 시사점을 주는 조항들이 많이 포함되었다. 그 중 중요한 것 하나는 개인정보의 수집·이용 및 제공에 있어 정보주체의 추가적인 동의를 필요로 하지 않는

08 개인정보 보호법 제2조(정의) 이 법에서 사용하는 용어의 뜻은 다음과 같다. 1. "개인정보"란 살아 있는 개인에 관한 정보로서 성명, 주민등록번호 및 영상 등을 통하여 개인을 알아볼 수 있는 정보(해당 정보만으로는 특정 개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 알아볼 수 있는 것을 포함한다)를 말한다.

09 개정 개인정보 보호법 제2조(정의) 이 법에서 사용하는 용어의 뜻은 다음과 같다. 1. "개인정보"란 살아 있는 개인에 관한 정보로서 다음 각 목의 어느 하나에 해당하는 정보를 말한다. 가. 성명, 주민등록번호 및 영상 등을 통하여 개인을 알아볼 수 있는 정보 나. 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.

10 개정 개인정보 보호법 제2조(정의) 이 법에서 사용하는 용어의 뜻은 다음과 같다. 1. "개인정보"란 살아 있는 개인에 관한 정보로서 다음 각 목의 어느 하나에 해당하는 정보를 말한다. 다. 가목 또는 나목을 제1호의2에 따라 가명처리함으로써 원래의 상태로 복원하기 위한 추가 정보의 사용 · 결합 없이는 특정 개인을 알아볼 수 없는 정보(이하 "가명정보"라 한다) 1의2. "가명처리"란 개인정보의 일부를 삭제하거나 일부 또는 전부를 대체하는 등의 방법으로 추가 정보가 없이는 특정 개인을 알아볼 수 없도록 처리하는 것을 말한다.

11 개정 개인정보 보호법은 '추가 정보'라는 표현을 사용한다. 법상 추가 정보의 개념이 무엇인지 명확하지는 않다. 이는 가명처리의 과정에서 생성될 수 있는 매칭테이블 뿐만 아니라 가명처리 과정에 적용될 수 있는 해시함수 등 가명화 기술까지 포함한다고 해석함이 합리적일 것이다. 추가 정보는, 공격자 등 개인정보처리자 이외의 제3자가 보유하는 "보조정보" 내지 "결합정보"와는 구별되는 개념인 것으로 생각할 수도 있다. 부가적인 정보라는 점에서는 마찬가지이지만, 정보의 보유주체나 원천이 크게 다르기 때문이다. 결국 법문상 표현되어 있는 추가 정보라는 표현의 해석에 있어, 주로 '가명처리 비밀' 등을 가리키는 것인지, 좀 더 넓게 보조정보 내지 결합정보를 아우르는 것인지에 대해 유의해야 한다. 현재로는 그 구체적인 의미는 명확하지 않다.

12 이 경우 쉽게 결합할 수 있는지 여부는 다른 정보의 입수 가능성 등 개인을 알아보는 데 소요되는 시간, 비용, 기술 등을 합리적으로 고려하여야 한다.

13 개정 개인정보 보호법 제58조의2(적용제외) 이 법은 시간 · 비용 · 기술 등을 합리적으로 고려할 때 다른 정보를 사용하여도 더 이상 개인을 알아볼 수 없는 정보에는 적용하지 아니한다.

양립가능성(compatibility)의 예외를 신설한 것이고,[14] 또 다른 하나는 통계작성, 과학적 연구, 공익적 기록보존 등의 목적을 위하여 정보주체의 동의 없이 가명정보를 처리하는 것을 허용한 것이다(개정 개인정보 보호법 제28조의2). 이하에서는 가명정보와 관련된 내용에 관하여 좀 더 상세하게 살펴본다.

개인정보처리자는 가명정보를 과학적 연구 등의 제한된 목적을 위하여 이용할 수 있고, 또한 제3자에게 제공할 수도 있다. 개인정보처리자가 제3자에게 가명정보를 제공하는 경우 특정 개인을 알아보기 위하여 사용할 수 있는 정보를 포함해서는 아니된다(개정 개인정보 보호법 제28조의2). 서로 다른 개인정보처리자들이 통계작성, 과학적 연구, 공익적 기록보존 등의 목적을 위하여 가명정보의 결합을 하는 것도 가능하다. 데이터 결합을 하는 경우에는 법에서 정한 '전문기관'이 해당 작업을 수행해야 한다(개정 개인정보 보호법 제28조의3 제1항).

개정된 데이터 3법에 따르면 가명처리된 정보에는 정보주체의 권리행사가 일정부분 제한된다. 구체적으로는, <표1>에 정리된 것과 같이 법에 명시된 정보주체의 권리에 대해 제한을 두게 된다.

개인정보 보호법 (제28조의7)	정보주체 이외로부터 수집한 개인정보의 수집출처 등 고지(제20조)
	개인정보의 파기(제21조)
	영업양도 등에 따른 개인정보의 이전 제한(제27조)
	개인정보 유출 통지(제34조 제1항)
	개인정보의 열람(제35조)
	개인정보의 정정·삭제(제36조)
	개인정보의 처리정지 등(제37조)
	개인정보의 수집·이용 동의 등에 대한 특례(제39조의3)
	개인정보 유출등의 통지·신고에 대한 특례(제39조의4)
	개인정보의 파기에 대한 특례(제39조의6)
	이용자의 권리 등에 대한 특례(제39조의7)
	개인정보 이용내역의 통지(제39조의8)
신용정보의 이용 및 보호에 관한 법률 (제40조의3)	개인신용정보의 제공 · 활용에 대한 동의(제32조 제7항)
	개인신용정보의 전송요구(제33조의2)
	신용정보 이용 및 제공사실의 조회(제35조)
	개인신용평점 하락 가능성 등에 대한 설명의무(제35조의2)
	신용정보제공·이용자의 사전통지(제35조의3)
	상거래 거절 근거 신용정보의 고지 등(제36조)
	자동화평가 결과에 대한 설명 및 이의제기 등(제36조의2)
	개인신용정보 제공 동의 철회권 등(제37조)
	신용정보의 열람 및 정정청구 등(제38조)
	신용조회사실의 통지 요청(제38조의2)
	개인신용정보의 삭제 요구(제38조의3)
	무료 열람권(제39조)
	채권자변동정보의 열람 등(제39조의2)
	신용정보주체의 권리행사 방법 및 절차(제39조의3)
	개인신용정보 누설통지 등(제39조의4)

<표1> 가명정보 적용제외 규정

다. 데이터 3법의 가명정보 관련규정

개정된 데이터 3법 중 개인정보 보호법과 신용정보법에 각각 가명정보에 관한 규정을 두고 있다. 두 법의 규정이 유사하기는 하지만 세부적인 차이가 나기도 한다. <표2>는 두 법에서 규정하고 있는 가명정보 관련 조항을 비교하여 정리한 것이다.

14 개정 개인정보 보호법 제15조 ③ 개인정보처리자는 당초 수집 목적과 합리적으로 관련된 범위에서 정보주체에게 불이익이 발생하는지 여부, 암호화 등 안전성 확보에 필요한 조치를 하였는지 여부 등을 고려하여 대통령령으로 정하는 바에 따라 정보주체의 동의 없이 개인정보를 이용할 수 있다. 개정 개인정보 보호법 제17조 ④ 개인정보처리자는 당초 수집 목적과 합리적으로 관련된 범위에서 정보주체에게 불이익이 발생하는지 여부, 암호화 등 안전성 확보에 필요한 조치를 하였는지 여부 등을 고려하여 대통령령으로 정하는 바에 따라 정보주체의 동의 없이 개인정보를 제공할 수 있다. 신용정보의 경우에도 양립가능성의 예외가 규정되어 있다(개정 신용정보의 이용 및 보호에 관한 법률 제32조 제6항 제9의3호).

	개인정보 보호법	신용정보법
연구 목적 등을 위한 처리	개인정보처리자는 통계작성, 과학적 연구, 공익적 기록보존 등을 위하여 정보주체의 동의 없이 가명정보를 처리할 수 있다(제28조의2 제1항).	신용정보회사[15] 등이 통계작성, 연구, 공익적 기록보존 등을 위하여 가명정보를 제공하는 경우 정보주체의 동의가 필요하지 않다. 이 경우 통계작성에는 시장조사 등 상업적 목적의 통계작성을 포함하며, 연구에는 산업적 연구를 포함한다(제32조 제6항 제9의2호)
	개인정보처리자가 위의 예외규정에 따라 가명정보를 제3자에게 제공하는 경우에는 특정 개인을 알아보기 위하여 사용될 수 있는 정보를 포함해서는 아니 된다(제28조의2 제2항)	
결합	통계작성, 과학적 연구, 공익적 기록보존 등을 위한 서로 다른 개인정보처리자 간의 가명정보의 결합은 보호위원회 또는 관계 중앙행정기관의 장이 지정하는 전문기관이 수행한다(제28조의3 제1항)	신용정보회사등[16]이 자기가 보유한 정보집합물을 제3자가 보유한 정보집합물과 결합하려는 경우에는 지정된 데이터 전문기관을 통하여 결합하여야 한다(제17조의2 제1항)
결합된 정보의 반출	결합을 수행한 기관 외부로 결합된 정보를 반출하려는 개인정보처리자는 가명정보 또는 시간·비용·기술 등을 합리적으로 고려할 때 다른 정보를 사용하여도 더 이상 개인을 알아볼 수 없는 정보에 해당하는 정보로 처리한 뒤 전문기관의 장의 승인을 받아야 한다(제28조의3 제2항)	데이터 전문기관이 결합된 정보집합물을 해당 신용정보회사등 또는 그 제3자에게 전달하는 경우에는 가명처리 또는 익명처리가 된 상태로 전달하여야 한다(제17조의2 제2항)
결합 등의 절차 · 방법	가명정보의 결합 절차와 방법, 전문기관의 지정과 지정취소 기준·절차, 관리·감독 뿐만 아니라 그 반출 및 승인 기준·절차 등 필요한 사항은 대통령령[17, 18, 19]으로 정한다(제28조의3 제3항)	정보집합물의 결합·제공·보관의 절차 및 방법은 대통령령[20]으로 정한다(제17조의2 제3항)
안전성 확보조치	개인정보처리자는 가명정보를 처리하는 경우에는 원래의 상태로 복원하기 위한 추가 정보를 별도로 분리하여 보관·관리하는 등 해당 정보가 분실·도난·유출·위조·변조 또는 훼손되지 않도록 대통령령[21]으로 정하는 바에 따라 안전성 확보에 필요한 기술적·관리적 및 물리적 조치를 하여야 한다(제28조의4 제1항)	신용정보회사등은 가명처리에 사용한 추가 정보를 대통령령으로 정하는 방법으로 분리하여 보관하거나 삭제하여야 한다(제40조의2 제1항).
		신용정보회사등은 가명처리한 개인신용정보에 대하여 제3자의 불법적인 접근, 입력된 정보의 변경·훼손 및 파괴, 그 밖의 위험으로부터 가명정보를 보호하기 위하여 내부관리계획을 수립하고 접속기록을 보관하는 등 대통령령으로 정하는 바에 따라 기술적·물리적·관리적 보안대책을 수립·시행하여야 한다(제40조의2 제2항)
가명정보 처리의 처리	누구든지 특정 개인을 알아보기 위한 목적으로 가명정보를 처리해서는 아니 되며(제28조의5 제1항), 개인정보처리자는 가명정보를 처리하는 과정에서 특정 개인을 알아볼 수 있는 정보가 생성된 경우에는 즉시 해당 정보의 처리를 중지하고, 지체 없이 회수 · 파기하여야 한다(제28조의5 제2항)	신용정보회사등은 영리 또는 부정한 목적으로 특정 개인을 알아볼 수 있게 가명정보를 처리하여서는 아니 된다(제40조의2 제6항).
		신용정보회사등은 가명정보를 이용하는 과정에서 특정 개인을 알아볼 수 있게 된 경우 즉시 그 가명정보를 회수하여 처리를 중지하고, 특정 개인을 알아볼 수 있게 된 정보는 즉시 삭제하여야 한다(제40조의2 제7항).

<표2> 개인정보 보호법과 신용정보법의 가명정보 관련규정 (1/2)

15 신용정보법 제9호의3을 적용하는 경우에는 데이터전문기관을 포함한다.

16 대통령령으로 정하는 자는 제외한다.

17 개인정보 보호법 시행령 일부개정령안(2020. 3. 31. 입법예고) 제29조의2(개인정보처리자 간 가명정보의 결합 등) ① 법 제28조의3제1항에 따른 전문기관(이하 “결합전문기관”이라 한다)에 가명정보의 결합을 신청하려는 개인정보처리자(이하 “결합신청자”라 한다)는 보호위원회가 정하여 고시하는 결합신청서를 해당 결합전문기관에 제출하여야 한다. ② 결합전문기관은 특정 개인을 알아볼 수 없도록 보호위원회가 정하여 고시하는 절차와 방법에 따라 가명정보를 결합하여야 한다. 이 경우 보호위원회는 결합전문기관이 특정 개인을 알아볼 수 없도록 하는데 필요한 지원업무를 한국인터넷진흥원이 수행하도록 할 수 있다. ③ 결합신청자는 보호위원회가 정하여 고시하는 바에 따라 결합전문기관에 설치된 안전성 확보에 필요한 기술적·관리적·물리적 조치가 된 공간(이하 “분석공간”이라 한다)에서 제2항에 따라 결합된 정보를 분석할 수 있다. ④ 제3항에도 불구하고 분석공간에서는 결합 목적을 달성하기 어렵거나 분석공간의 이용이 어려운 경우로서 결합신청자가 제2항에 따라 결합된 정보의 반출을 신청하는 경우, 결합전문기관은 개인을 다시 알아볼 가능성 등을 고려하여 보호위원회가 정하여 고시하는 바에 따라 평가한 후 반출을 승인할 수 있다. ⑤ 결합전문기관은 이 조에 따른 결합, 반출 등에 필요한 비용을 결합신청자에게 청구할 수 있다. ⑥ 제1항부터 제5항까지 규정한 사항 외에 가명정보 결합 절차와 방법, 반출 및 승인 등에 필요한 세부사항은 보호위원회가 정하여 고시한다.

18 개인정보 보호법 시행령 일부개정령안(2020. 3. 31. 입법예고) 제29조의4 (관리·감독 등) ① 결합전문기관을 지정한 보호위원회 또는 관계 중앙행정기관의 장은 해당 결합전문기관에 대하여 결합전문기관으로서의 업무 수행능력 및 기술·시설 유지 여부 등에 대한 관리·감독을 하여야 한다. ② 보호위원회는 결합전문기관에 대한 가명정보의 결합 및 반출 승인 과정에서의 법 위반 여부, 결합 신청자에 대한 가명정보 처리 실태, 그 밖에 가명정보의 안전한 처리를 위하여 필요한 사항으로서 보호위원회가 정하여 고시하는 사항에 대한 관리·감독을 하여야 한다.

19 개인정보 보호법 시행령 일부개정령안(2020. 3. 31. 입법예고) 제29조의3(결합전문기관의 지정 및 지정취소) ① 보호위원회 또는 관계 중앙행정기관의 장은 법 제28조의3제1항에 따라 보호위원회가 정하여 고시하는 기준을 갖춘 법인, 단체 또는 기관을 결합전문기관으로 지정할 수 있다. 이 경우 보호위원회는 가명정보의 결합, 결합된 정보의 처리 지원 및 반출심사를 안정적으로 수행하는데 필요한 인력·조직, 시설·장비, 재정 능력 등을 고려하여 지정 기준을 정하여야 한다. ② 결합전문기관 지정의 유효기간은 지정 받은 날부터 3년으로 하며, 재지정할 수 있다. ③ 보호위원회 또는 관계 중앙행정기관의 장은 결합전문기관이 다음 각 호의 어느 하나에 해당하는 경우에는 전문기관의 지정을 취소할 수 있다. 다만, 제1호 또는 제2호에 해당하는 경우에는 지정을 취소하여야 한다. 1. 거짓이나 부정한 방법으로 결합전문기관으로 지정을 받은 경우 2. 결합전문기관 스스로 지정 취소를 요청하거나 폐업한 경우 3. 제1항에 따른 결합전문기관으로 지정되기 위한 기준을 충족하지 못하게 된 경우 4. 결합 및 분석 등과 관련된 정보의 유출 등 침해사고가 발생한 경우 5. 그 밖에 법 또는 이 영에 따른 의무를 위반한 경우 ④ 보호위원회 또는 관계 중앙행정기관의 장은 제3항에 따라 결합전문기관의 지정을 취소하려는 경우에는 청문을 하여야 한다. ⑤ 보호위원회 또는 관계 중앙행정기관의 장은 전문기관을 지정·재지정·지정 취소하였을 때에는 이를 고시하여야 한다. 이 경우, 관계 중앙행정기관의 장이 전문기관을 지정·재지정·지정 취소한 경우에는 보호위원회에 통보하여야 한다. ⑥ 제1항에 따른 결합전문기관 지정 절차, 세부 지정기준, 심사방법, 제2항에 따른 재지정 등에 관하여 필요한 사항은 보호위원회가 정하여 고시한다.

20 신용정보의 이용 및 보호에 관한 법률 시행령(2020. 3. 31. 입법예고) 제14조의2(정보집합물의 결합 등) ① 법 제17조의2제1항에서 “대통령령으로 정하는 자”란 법 제45조의3에 따른 상거래기업 및 법인(이하 “상거래 기업 및 법인”이라 한다)을 말한다. ② 법 제17조의2제1항에 따라 정보집합물을 결합하려는 신용정보회사등과 제3자(이하 이 조에서 “결합의뢰기관”이라 한다)는 공동으로 데이터전문기관에 금융위원회가 정하여 고시하는 양식에 따라 정보집합물의 결합을 신청하여야 한다. ③ 결합의뢰기관 및 데이터전문기관은 정보집합물을 결합 · 제공 ·

	개인정보 보호법	신용정보법
심사 · 추정		신용정보회사등은 개인신용정보에 대한 익명처리가 적정하게 이루어졌는지 여부에 대하여 금융위원회에 그 심사를 요청할 수 있다(제40조의2 제3항).
		금융위원회가 위의 요청에 따라 심사하여 적정하게 익명처리가 이루어졌다고 인정한 경우 더 이상 해당 개인인 신용정보주체를 알아볼 수 없는 정보로 추정한다(제40조의2 제4항).
		금융위원회는 위의 심사 업무(제40조의2 제3항) 및 인정 업무(제40조의2 제4항)에 대해서는 대통령령으로 정하는 바에 따라 데이터 전문기관에 위탁할 수 있다.
기록보존	개인정보처리자는 가명정보를 처리하고자 하는 경우에는 가명정보의 처리 목적, 제3자 제공 시 제공받는 자 등 가명정보의 처리 내용을 관리하기 위하여 대통령령으로 정하는 사항에 대한 관련 기록을 작성하여 보관하여야 한다(제28조의4 제2항)	신용정보회사등은 개인신용정보를 가명처리나 익명처리를 한 경우 개인신용정보를 가명처리한 경우에는 가명처리한 날짜와 정보의 항목 및 사유와 근거를, 개인신용정보를 익명처리한 경우에는 익명처리한 날짜와 그 항목 및 사유와 근거를 3년간 보존하여야 한다(제40조의2 제8항)
위반시 제재	개인정보처리자가 특정 개인을 알아보기 위한 목적으로 가명정보를 처리한 경우 전체 매출액의 100분의 3 이하에 해당하는 금액을 과징금으로 부과할 수 있으며, 매출액이 없거나 매출액의 산정이 곤란한 경우로서 대통령령으로 정하는 경우에는 4억원 또는 자본금의 100분의 3 중 큰 금액 이하로 과징금을 부과할 수 있다(제28조의6 제1항)[22].	제17조의2제1항을 위반하여 정보집합물을 결합한 자는 5년 이하의 징역 또는 5천만원 이하의 벌금에 처한다(제50조 제2항 제4의2호)
	제28조의3을 위반하여 가명정보를 처리하거나 제3자에게 제공한 자 및 그 사정을 알면서도 영리 또는 부정한 목적으로 가명정보를 제공받은 자와 제28조의5제1항을 위반하여 특정 개인을 알아보기 위한 목적으로 가명정보를 처리한 자는 5년 이하의 징역 또는 5천만원 이하의 벌금에 처한다(제71조 제4의2호, 제4의3호).	법인의 대표자나 법인 또는 개인의 대리인, 사용인, 그 밖의 종업원이 그 법인 또는 개인의 업무에 관하여 제50조의 위반행위를 하면 그 행위자를 벌하는 외에 그 법인 또는 개인에게도 해당 조문의 벌금형을 과(科)한다. 다만, 법인 또는 개인이 그 위반행위를 방지하기 위하여 해당 업무에 관하여 상당한 주의와 감독을 게을리하지 아니한 경우에는 그러하지 아니하다(제51조).
	제28조의4 제1항을 위반하여 안전성 확보에 필요한 조치를 하지 아니하여 개인정보를 분실·도난·유출·위조·변조 또는 훼손당한 자는 2년 이하의 징역 또는 2천만원 이하의 벌금에 처한다(제73조 제1호).	제17조의2 제2항을 위반하여 가명처리 또는 익명처리가 되지 아니한 상태로 전달한 자에게는 5천만원 이하의 과태료를 부과한다(제52조 제2항 제2의2호).
	제28조의5 제2항을 위반하여 개인을 알아볼 수 있는 정보가 생성되었음에도 이용을 중지하지 아니하거나 이를 회수·파기하지 아니한 자에게는 5천만원 이하의 과태료를 부과한다(제7조 제2항 7의2호) 제28조의4 제2항을 위반하여 관련 기록을 작성하여 보관하지 아니한 자는 1천만원 이하의 과태료를 부과한다(제75조 제4항 제6의2호)	가명정보의 처리와 안전성 확보에 관한 규정을 위반하여 가명처리에 사용한 추가 정보를 분리하여 보관하거나 삭제하지 아니한 자, 가명처리한 개인신용정보에 대하여 기술적·물리적·관리적 보안대책을 수립·시행하지 아니한 자, 처리를 중지하거나 정보를 즉시 삭제하지 아니한 자에게는 3천만 원 이하의 과태료를 부과한다(제52조 제3항 제16호 내지 제18호).

<표2> 개인정보 보호법과 신용정보법의 가명정보 관련규정 (2/2)

보관하는 경우에는 다음 각 호의 사항을 모두 준수하여야 한다. 1. 결합의뢰기관이 정보집합물을 데이터전문기관에 제공하는 경우 다음 각 목의 조치를 하여 제공할 것 가. 하나의 정보집합물과 다른 정보집합물간에 둘 이상의 정보를 연계, 연동하기 위하여 사용되는 정보는 해당 개인을 식별할 수 없으나 구별할 수 있는 정보(이하 “결합키”라 한다)로 대체할 것 나. 개인신용정보가 포함된 정보집합물은 가명처리할 것 2. 결합의뢰기관이 결합키를 생성하는 절차와 방식은 금융위원회가 정하여 고시하는 바에 따라 결합의뢰기관간 상호 협의하여 결정할 것 3. 결합의뢰기관이 데이터전문기관에 정보집합물을 제공하거나 데이터전문기관이 결합한 정보집합물을 결합의뢰기관에 전달하는 경우에는 해당 정보집합물의 내용을 제3자가 알 수 없도록 암호화 등의 보호조치를 하여 전달할 것 4. 데이터전문기관은 결합된 정보집합물을 결합의뢰기관에 전달하기 전 결합키를 삭제하거나 금융위원회가 정하여 고시하는 방법으로 대체할 것 5. 데이터전문기관은 결합된 정보집합물의 가명처리 또는 익명처리의 적정성을 평가한 후 적정하지 않다고 판단되는 경우 다시 가명처리 또는 익명처리하여 전달할 것 6. 데이터전문기관은 결합한 정보집합물을 결합의뢰기관에 전달한 후 결합한 정보집합물 및 결합전 정보집합물을 지체없이 삭제할 것 ④ 데이터전문기관은 금융위원회가 정하여 고시하는 방법에 따라 결합관련 사항을 기록 · 관리하고 1년에 1회 정기적으로 금융위원회에 보고하여야 한다. ⑤ 데이터전문기관은 데이터 결합 등에 필요한 비용을 결합의뢰기관에 청구할 수 있다. ⑥ 그 밖에 정보집합물 결합 · 제공 · 처리 · 보관의 절차 및 방법과 관련하여 필요한 사항은 금융위원회가 정하여 고시한다.

21 개인정보 보호법 시행령 일부개정령안(2020. 3. 31. 입법예고) 제29조의5(가명정보 등의 안전성 확보조치 등) ① 개인정보처리자는 법 제28조의4제1항에 따라 가명정보 및 추가 정보에 대하여 다음 각 호의 안전성 확보에 필요한 조치를 하여야 한다. 1. 내부 관리 계획의 수립·시행 등 제30조에 따른 개인정보의 안전성 확보 조치(이 경우 "법 제29조"는 "법 제28조의4제1항"으로, "개인정보"는 "가명정보 및 추가 정보"로 보며, 법 제28조의3제2항에 따라 반출한 가명정보를 포함한다) 2. 추가 정보의 별도 분리 보관 및 추가 정보에 대한 접근 권한 분리 3. 가명정보 또는 추가 정보에 대한 접근 권한 관리 및 물리적·기술적 안전조치에 관한 내부 관리 계획 수립·시행 ② 개인정보처리자가 가명정보를 처리하고자 하는 경우 가명정보의 처리 목적, 처리 및 보유기간, 추가 정보의 이용 및 파기 등 보호위원회가 정하여 고시하는 사항을 작성하여 보관하여야 한다. ③ 개인정보처리자는 가명정보의 처리 목적이 달성되거나 가명정보 보유 기간이 경과한 때에는 그 가명정보를 지체 없이 파기하여야 한다. 이 경우, 해당 가명정보는 제16조제1항제1호·제2호의 구분에 따른 방법으로 파기하여야 한다.

22 개인정보 보호법 시행령 일부개정령안(2020. 3. 31. 입법예고) 제29조의6(가명정보 처리에 대한 과징금의 부과기준 등) ① 법 제28조의6에서 "전체 매출액"이란 해당 개인정보처리자의 직전 3개 사업연도의 연평균 매출액을 말한다. 다만, 해당 사업연도 초일 현재 사업을 개시한지 3년이 되지 아니하는 경우에는 그 사업개시 후 직전 사업연도 말일까지의 매출액을 연평균 매출액으로 환산한 금액을 말하며, 해당 사업연도에 사업을 개시한 경우에는 사업개시일부터 위반행위일까지의 매출액을 연매출액으로 환산한 금액을 말한다. ② 법 제28조의6에서 "대통령령으로 정하는 경우"란 다음 각 호의 어느 하나에 해당하는 경우를 말한다. 1. 영업을 개시하지 아니하거나 영업을 중단하는 등의 사유로 영업실적이 없는 경우 2. 재해 등으로 인하여 매출액 산정자료가 소멸되거나 훼손되는 등 객관적인 매출액의 산정이 곤란한 경우 ③ 보호위원회는 제1항에 따른 매출액 산정을 위하여 재무제표 등 자료가 필요한 경우 20일 이내의 기간을 정하여 해당 개인정보처리자에게 관련 자료의 제출을 요청할 수 있다. ④ 법 제28조의6에 따른 과징금의 산정기준과 산정절차는 별표 1의3과 같다.

라. 가명정보와 가명처리의 개념에 관련된 쟁점들

1) 가명처리의 의미

가명처리는 개인정보의 일부를 삭제하거나 일부 또는 전부를 대체하는 등의 방법으로 추가 정보가 없이는 특정 개인을 알아볼 수 없도록 처리하는 것이다.[23] 고전적으로 가명(pseudonym)은 정보주체의 식별(identification)과 관계된 속성을 대체하여 표시하는 것을 의미하는 것으로 흔히 해석되는데,[24] "추가 정보가 없이는" 부분을 제외한다면 기존에 언급되던 비식별 조치(de-identification)의 한 형태인 것으로 볼 수도 있다. 그렇다면 가명처리가 비식별 조치를 의미하는 것인지 생각해 볼 필요가 있다. 비식별 조치의 의미를 규정한 "비식별조치 가이드라인"은 가명처리가 비식별 조치의 한 방법이라고 설명하였다.[25] 다만 이때 가명처리가 개인정보보호를 위해 이용한 유일한 방법이라면 가명처리를 통해 충분한 비식별 조치가 이루어진 것으로 보기 어렵다는 설명도 있다. 이렇게 보면, 가명처리는 비식별 조치의 한 방법인 것으로 이해할 수는 있지만, 구분하여 파악하는 것이 더 현실적이다.[26]

이 점에 비추어 개정 개인정보 보호법은 일반적인 의미의 익명정보와 식별정보 사이에 위치한 가명정보에 대해 그 개념을 규정하고, 가명정보에 대해서는 이용이나 제공을 부분적으로 허용한 것으로 볼 수 있다. 다른 한편, 기존의 비식별조치 가이드라인에는 속성정보의 일부를 치환하거나 임의의 잡음을 더하는 마스킹 기술과, k-익명성, l-다양성, t-근접성 등 신원정보의 일반화 기술에 관한 설명이 담겨있는데, 기존의 개인정보 보호법상으로는 이러한 기술이 비식별처리의 수단으로서 사용된 경우에만 정보주체의 동의의 범위를 초과한 이용과 제공이 가능하다.[27]

2) 추가 정보의 의미

식별 및 재식별에 관한 기존의 논의는, 비식별정보에 추가적 정보나 임의의 정보를 결합하여 정보주체의 (재)식별이 가능하도록 시도하는 방식의 공격(linkage attack)을 흔히 상정하고 전개되었다. 이 과정에서 다양한 관련 정보 - 특히 준식별자(quasi-identifier) 또는 속성정보(attributes) 등의 관련 정보 - 가 결합의 고리가 되기도 한다는 점이 확인되기도 하였다.[28] 여기서 "정보"는 정보주체의 식별에 도움이 될 수 있는 다양한 정보를 의미하는데, 이를 보조정보(auxiliary information), 외부지식(external knowledge), 배경지식(background knowledge) 등 여러 이름으로 부른다.[29] 이러한 정보는 주로 개인정보처리자 외부의 제3자가 보유하고 있을 수 있는 정보를 상정한다. 아래에서는 제3자가 보유할 수 있는 이러한 유형의 정보를 통칭하여 "보조정보"라 한다.

한편 가명처리는 암호화 키, 매핑표(mapping table) 등 일정한 변수(parameter)를 설정하여 데이터를 변환하는 과정을 요구하는데, 이 과정에서 이용되는 암호화 키나 매핑표 등을 가명처리 비밀(pseudonymization secret)이라고 부를 수 있다.[30] 그런데 가명처리 비밀은 가명처리의 과정에서 활용되는 것은 물론 그 반대로 복원의 열쇠로도 작용할 수 있다. 가명처리 비밀이

23 국제표준화기구(International Organization for Standardization, 이하 'ISO') 문서에서는, 정보주체와 개인정보의 연결을 삭제하면서, 정보주체와 관계된 일정한 속성과 가명의 연결을 더하는 일종의 비식별처리("particular type of de-identification that both removes the association with a data subject and adds an association between a particular set of characteristics relating to the data subject and one or more pseudonyms")라고 정의하였다(ISO 25237:2017).

24 "비식별 조치"는 "정보집합물(데이터 셋)에서 개인을 식별할 수 있는 요소를 전부 또는 일부 삭제하거나 대체하는 등의 방법을 활용, 개인을 알아볼 수 없도록 하는 조치"인 것으로 설명될 수 있다. 국무조정실 외, "개인정보 비식별 조치 가이드라인" (2016) 참조.

25 ISO 문서에서도 가명처리가 일종의 비식별처리("particular type of de-identification")라고 정의하였다.

26 정보주체의 식별이 가능한 추가적 정보의 존재가 가명처리의 고유한 성격이라고 보아야 하는가? 이를 위해서는 추가적 정보(또는 법에서 정한 "추가 정보")의 의미가 우선적으로 확정되어야 하므로, 이에 관해서는 다음의 절에서 논의하도록 한다.

27 반드시 익명처리의 수단이 아니라 하여도, 비식별처리 기술을 적용하는 것은 일반적으로 개인정보 보호의 수준을 높이는 효과를 가져온다. 따라서, 예컨대, 마스킹 기술은 일반적으로는 정보주체의 입장에서도 유익한 것이다. 개정 개인정보 보호법 제3조 제7항은 익명처리와 가명처리의 원칙을 선언하는데, 동일한 의도로 이해할 수 있다.

28 L. Sweeney. k-anonymity: a model for protecting privacy. International Journal on Uncertainty, Fuzziness and Knowledge-based Systems, 10 (5), (2002). ; Arvind Narayanan and Vitaly Shmatikov, Robust De-anonymization of Large Sparse Datasets, IEEE Symposium on Security and Privacy (2008). ; Ohm, Paul, Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization, UCLA Law Review, Vol. 57, p. 1701 (August 13, 2009).

29 참고로, 보조정보는 공격자에게 주어진 지식이므로, 익명처리가 일정한 수준의 유의한 프라이버시 보호를 사전적으로 보장하기는 어렵다(Cynthia Dwork Moni Naor Toniann Pitassi Guy Rothblum, "Differential Privacy Under Continual Observation", STOC '10: Proceedings of the 42nd ACM symposium on Theory of computing (June 2010)). 익명처리는 불가능하고, 따라서, 정보주체의 동의의 내용과 범위를 초과한 개인정보의 이용은 절대적으로 금지되어야 한다는 주장은 보조정보의 위험을 적극적으로 고려한 것이다.

30 매핑표 등 정보주체의 일정한 신원정보와 가명정보가 열거된 형식의 가명처리 비밀은 (개인정보의 적법한 이용의 방식으로서) 생성된 식별정보에 해당한다고 해석할 수 있다. 이 비밀의 이용은 개정 개인정보 보호법 제15조 제1항, 제공은 개정 개인정보 보호법 제28조의2 제2항, 안전조치의무는 개정 개인정보 보호법 제28조의4 제1항 등 여러 조항의 적용을 받는다.

공격자에게 주어진다면 권한이 없는 자에 의한 정보주체의 식별이 가능해질 수도 있다. 이를 고려하여, 가명처리 비밀 등의 추가 정보는 별도로 분리보관하게 하고 추가 정보에 대한 접근권한을 분리하도록 함으로써 가명처리 비밀을 제3자에게 제공하지 못하도록 정하고 있다(개정 개인정보 보호법 제28조의2 제2항, 제28조의4 제1항, 시행령 개정령안 제29조의5 제1항 제2호).

이처럼 보조정보와 가명처리 비밀은 식별 또는 재식별의 위험을 유발하는 주요 원인이 될 수 있다. 이 점에 착안하여 개인정보 보호법은 개인정보의 의미를 식별의 가능성과 연관시킨다.[31] 시간·비용·기술 등을 합리적으로 고려하였을 때 다른 정보를 사용하여 개인을 알아볼 수 있는 정보가 개인정보라고 규정한 것이다(개인정보 보호법 제58조의2 반대해석).[32] 식별이 합리적으로 가능하도록 만드는 정보가 존재한다는 점에서 가명정보는 개인정보이다. 그런데 원래의 상태로 복원하기 위한 추가 정보의 사용과 결합이 없으면 가명정보의 식별은 가능하지 않은 것으로 해석된다. 달리 말하면, 이러한 추가 정보는 가명정보의 식별을 가능하게 해주는 정보다.

이때 개정 개인정보 보호법에서 정한 "추가 정보"와 "다른 정보"의 정확한 의미가 무엇인지, 서로 어떻게 구분되는지 문제될 수 있다. 쉽게 생각해 볼 수 있는 구분으로, 전자는 가명처리 비밀을, 후자는 보조정보를 의미한다고 해석할 수 있는데, 명확하지 않다. 또한, 후자는 논리적으로 전자를 포함하여 이해하는 경우도 있는데, 개정 개인정보 보호법이 그렇게 개념 설정을 하고 있는 것인지 여부도 명확하지 않다. 다른 한편, "추가 정보"가 보조정보를 포함한다고 해석하는 것에도 무리가 있을 수 있다.[33] 이는 보조정보를 활용하는 방식으로는 가명정보를 통한 재식별이 불가능하고 프라이버시 침해의 위험도 없어야 한다는 것인지에 대한 질문으로 귀결된다.

이 질문은 개념적으로나 실무적으로 매우 중요한 질문이다. 그 해답에 따라, 가명정보와 익명정보 사이에 프라이버시 보호의 수준이 역전될 수도 있는 이론적 가능성이 있기 때문이다. 개념상 가명정보는 처리의 목적과 방식이 제한된 특수한 유형의 개인정보인 반면(개정 개인정보 보호법 제28조의2 제1항, 제2항,

31 구 공공기관의개인정보보호에관한법률(1994. 1. 7. 법률 제4734호로 제정된 것) 이후의("개인정보"는 "생존하는 개인에 관한 정보로서 당해 정보에 포함되어 있는 성명·주민등록번호등의 사항에 의하여 당해 개인을 식별할 수 있는 정보(당해 情報만으로는 특정개인을 識別할 수 없더라도 다른 情報와 용이하게 結合하여 識別할 수 있는 것을 포함한다)"를 의미한다고 하였다.) 개인정보의 정의는, 합리적으로 고려하여야 한다는 기준을 제외한다면, 사실상 동일한 것이다.

32 그런데, 개정된 개인정보 보호법 제58조의2 조항은, '다른 정보를 사용하여도 더 이상 개인을 알아볼 수 없는 정보에는 적용하지 아니한다'고 규정하여 정보의 "사용"만 식별의 방식으로서 포함하였다. 그렇다고 하여 결합을 의도적으로 제외한 것으로 보기는 어렵다.

33 개정된 개인정보 보호법 제28조의4 제1항, 시행령 개정령안 제29조의5 제1항 제2호 등을 합목적적으로 해석하면 그와 같이 해석하기에 한계가 있을 수 있기 때문이다.

제28조의3, 제28조의4), 익명정보는 개인정보가 아니므로(개정 개인정보 보호법 제58조의2) 임의의 목적과 방식으로서 처리할 수 있다. 이를 고려하면, 주어진 정보가 익명정보인지 여부에 대한 판단은 가명정보인지 여부에 대해 판단하는 경우보다 상대적으로 더욱 엄격한 기준이 적용되어야 할 것이다. 그런데 개인정보 보호법은 익명정보에는 합리적 기준의 심사를 명시적으로 규정하고 있는 한편, 가명정보에는 어떤 기준의 심사를 할 것인지 명시적으로 정하고 있지 않다. 그런데 만일 가명정보에 대한 판단에 엄격한 기준을 적용하여, 가령 추가 정보 없이는 재식별 가능성이 전혀 없을 것(zero-probability)이 요구된다면, 추가 정보의 개념을 넓게 해석하지 않는 한 이 기준은 현실적으로는 충족이 거의 불가능할 것이다. 왜냐하면, 보조정보는 공격자에게 주어진 정보이므로 이에 관해 사전적으로 예상하는 데에 한계가 있을 수밖에 없기 때문이다.[34] 한편 보조정보를 이용한 공격의 가능성을 모두 고려하여 가명처리를 해야 한다면, 이는 주어진 데이터베이스 안에서 식별자(identifier)나 준식별자(quasi-identifier)는 물론 그 이외의 온갖 속성정보(attributes)를 모두 고려해야 한다는 뜻이 된다. 그 경우 가명처리의 과정은 매우 복잡한 과정이 되고, 그에 대한 평가과정 또한 복잡해질 수밖에 없다.[35] 또한, 가명정보의 개념이 법에 명시적으로 도입된 것은, 속성정보가 대체로 유지되는 것을 전제로 데이터 분석의 활용도를 높이고자 하는 정책적 의도가 반영된 것일 가능성이 높은데, 보조정보를 통한 결합위험이 없어야 한다는 제약이 있다면 가명정보의 활용가능성 자체가 일반적으로 낮아질 수밖에 없을 것이다.

이와 별개로, 재식별 과정을 거쳐 "원래의 상태로 복원" 된다는 것의 의미가 무엇인지도 불분명하다. 이는, 특정 데이터베이스를 가정할 때, 데이터베이스 전체가 복원되는 경우만을 의미하는 것인지 여부에 관한 질문이다. 예를 들어, 원본 데이터베이스 전체가 복원되지 않더라도 해당 데이터베이스에 담긴 신원정보가 일부 개인에 대해서 복원이 된다면 이를 원래의 상태로 복원이 된 것으로 인정될 것인지 문제될 수 있다. 나아가, 공격을 통한 복원의 대상이

34 기존의 연구는 전문가 그룹의 (비식별처리) 적정성 심사가 필요하다고 주장하기도 하였다. 예컨대, 의료정보의 비식별처리를 위해 윤리위원회(ethics board) 심사와 제한적 조건의 설정(예컨대, 익명이 유효한 기간은 일반적으로 18~24개월)을 정해 놓는 방식이 제안되기도 한다(El Emam K, Rodgers S, Malin B. "Anonymising and sharing individual patient data.", PMC, (2015)).

35 EU의 개인정보보호일반법인 GDPR(General Data Protection Regulation)에 규정된 추가적 정보("additional information")의 경우에도, 이를 가명처리 비밀을 의미하는 것으로 좁혀서 해석할 수도 있고 다양한 보조정보를 의미하는 것으로 넓게 해석할 수도 있다. 다음과 같이 넓게 해석한 문헌도 볼 수 있다. "If OSS has additional knowledge on a certain user's characteristics, and is trying to uncover that user's data records from the pseudonymised database it gets from SN, every piece of additional information may become critical." (ENISA, "Pseudonymisation Techniques and Best Practices", (2019)).

데이터베이스 전체가 아니라 데이터베이스에 담긴 특정한 개인에 관한 단편적인 정보라면, 공격자 입장에서는 부분적 복원으로도 목적달성이 될 수 있고 나아가 복원을 통한 완벽한 정확성 확보가 필요하지 않을 수도 있다. 가령 가명정보에 대한 재식별 공격의 결과, 해당 정보가 절반의 확률로 A 정보주체와 연결되고 절반의 확률로 B 정보주체와 연결된다면, 구체적인 맥락에 따라서는 그것만으로 상당한 프라이버시 침해가 가능할 수도 있기 때문이다. 확률론적인 복원이나 재식별의 가능성까지 고려한다면, 추가 정보는 매핑표의 일부분뿐만 아니라 확률적으로 계산하는 매핑 함수도 포함하여 넓게 고려해야 할 수도 있다.[36]

3) 가명정보의 처리와 결합

개인정보의 처리에는 정보주체의 동의가 필요한 것이 일반적인 원칙이지만 가명정보의 처리는 그 예외이다. 일반적으로 개인정보보호와 관련한 논의에 가명정보의 개념이 등장한 배경에는, 이를 통해 프라이버시 침해 위험을 줄이도록 하는 '안전장치(safeguards)'로서의 역할에 관한 논의가 있다. 가명처리가 이루어진 정보에 대해서는 일반적으로 재식별 가능성이 낮아지면서 프라이버시 침해 가능성도 함께 낮아질 것이기 때문이다. 그런데 이로부터 더 나아가, 데이터 3법을 통해 가명처리가 이루어지기만 하면 더 이상 프라이버시 침해 가능성을 별도로 고려할 필요 없이 정보를 활용할 수 있는 새로운 경로가 열린 것으로 해석될 수도 있다. 법에서 요구하는 기준을 충족하여 가명처리가 일단 진행되고 나면 그 이후에는 (재식별 금지 요건 이외에) 데이터에 대한 별도의 프라이버시 보호가 요구될 수 없는 것인지, 만일 요구될 수도 있다면 그 내용은 어떤 것인지는 명확하지 않다.

다른 한편, 가명정보의 결합은 가명정보 및 추가 정보의 전송 그리고 결합전문기관에서의 가명정보 결합이 포함된 과정으로서, 이를 통해 가명정보의 효용은 전반적으로 높아질 수 있는 한편 프라이버시의 침해위험도 높아지게 되는 것이 일반적이다. 관련된 전체의 과정 중에서 재식별 공격이 발생할 수 있는 상황도 늘어날 수 있을 뿐더러 결합 이후 더 풍부한 내용이 담기게 된 데이터에 대해 보조정보를 이용하여 재식별 공격이 벌어질 가능성도 있을 것이기 때문이다. 개정 개인정보 보호법의 시행령 개정령안에 따르면, 프라이버시 침해위험을 통제하기 위하여 분석은 지정된 공간에서[37] 진행하여야 함이 원칙이고,

36 더 나아가 재식별 공격에 따른 프라이버시 침해가능성의 맥락에서는, 개인정보의 복원이 공격의 유일한 목적이라고 보기도 어렵다. 예컨대, 공격자에게 가명처리된 데이터베이스가 유출되었고, 해당 데이터베이스에 담긴 구성원 상당수가 HIV 양성이라고 가정해보자. 그러한 경우에, 공격의 대상이 되는 정보주체가 해당 데이터베이스에 포함된 구성원인지 확인하는 것만으로도 프라이버시 침해가 발생할 수 있다.

37 결합전문기관에 설치된 안전성 확보에 필요한 기술적·관리적·물리적 조치가 이루어진 공간을 의미한다. 신용정보법 시행령 일부개정령안은 이와는 다른 방식을 규정하고 있다.

예외적으로만 결합된 가명정보를 반출할 수 있다(시행령 제29조의2 제3항 및 제4항).

4) 정보주체의 사생활 침해를 최소화하는 방법으로서의 가명처리

가명처리는 정보주체의 프라이버시를 보호해주는 주요한 도구로 작용할 수 있다. 개인정보 보호법은 정보주체의 사생활 침해를 최소화하는 처리의 방식을 사용하여야 한다는 원칙과 가명의 사용이 가능한 경우는 가명에 의하여 처리될 수 있도록 하여야 한다는 원칙을 선언하고 있다(개정 개인정보 보호법 제3조 제6항, 제7항)[38] 적절한 절차와 환경을 전제로 할 때, 가명처리된 개인정보는 프라이버시를 보호해 주는 동시에 과학적 연구 등을 위해 유용하게 활용될 수 있다. 속성정보가 대체로 유지되면서 이를 통해 추출해 낼 수 있는 효용(data utility)은 유지되는 한편, 가명처리를 통해 정보주체의 식별가능성이 낮아지는 효과가 동시에 나타날 수 있기 때문이다. 실제로 유럽연합의 Article 29 Data Protection Working Party는 신약개발을 위한 임상시험의 과정에서 일반적으로 키 코드가 사용된(key-coded) 가명정보를 이용한다고 하고, 가명처리가 과학적 연구의 현실과 긴밀한 관계를 가지고 있다고 소개한 바 있다.[39,40]

38 GDPR은 Art. 6(4)(e), 25(1), 32(1)(a), 40(2)(d) 조항과, Recital (28), (29), (78) 등으로 가명처리가 프라이버시 보호의 수단이라는 점을 명시적으로 규정하고 있다. 예를 들어, Recital (28) 규정 참조("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").

39 "This is particularly relevant in the context of research and statistics."; "This sort of data is commonly used in clinical trials with medicines. Directive 2001/20 of 4 April 2001 on the implementation of good clinical practice and the conduct of clinical trials lays down a legal framework for the pursuit of these activities."(Article 29 Data Protection Working Party, "Opinion 4/2007 on the concept of personal data" (2007)).

40 기존의 Directive 95/46/EC 내용("Whereas the further processing of personal data for historical, statistical or scientific purposes is not generally to be considered incompatible with the purposes for which the data have previously been collected provided that Member States furnish suitable safeguards")은, 회원국들이 적합한 안전성 조치(safeguard)를 규정한다면, 과학적 연구는 개인정보를 수집한 당시의 목적과 양립할 수 있다고 설명하였을 뿐, 가명처리의 의미를 규정하지는 아니하였다. 그리고 유럽연합집행위원회(European Commission)에서 제안한 GDPR의 초안도 Directive 95/46/EC와 동일한 방식으로서 개인정보와 익명정보의 의미만 규정하였다(European Commission, 2012). 그런데 Article 29 Data Protection Working Party가 가명처리의 의미를 명시적으로 규정하여야 한다는 의견("The Working Party believes that the concept of pseudonymisation should be introduced more explicitly in the instrument (for example by including a definition on pseudonymised data, consistent with the definition of personal data), as it can help to achieve better data protection, for example, in the context of data protection by design and default.")을 제시하였고, 가명처리의 의미를 명시적으로 규정한 의안이 통과되었다(Article 29 Data Protection Working Party, 2012).

이러한 가명처리를 통해 프라이버시가 보호되는 방식으로 개정 개인정보 보호법에 규정된 내용을 4단계로 나누어 파악할 수 있다. 첫째는 가명처리의 적정성 심사이며, 둘째는 가명처리된 개인정보를 이용·제공할 때에 요구되는 목적에 대한 심사이고, 셋째는 가명처리시 준수해야 하는 방법과 절차이며, 넷째는 가명처리시 요청되는 안전조치의무와 금지의무이다. 가명정보를 결합하는 경우에도 이러한 기준이나 원칙은 마찬가지로 적용된다.

가명처리 관련규정들을 해석하는 데에 있어, 실무적으로는 상당히 다양한 해석이나 입장이 나타날 수 있다. 가장 넓게는, 개인의 신원정보 등 식별자나 준식별자에 초점을 맞추는 것으로 충분할 것인지, 이와 다르게 폭넓고 다양한 가능성을 고려해야 할지에 관한 문제로 나타난다. ① 프라이버시 보호의 내용과 기준을 설정함에 있어 신원정보 유형의 정보에 대한 보호만으로 충분하다는 견해를 따르면 가명처리의 적정성 심사는 신원정보의 보호를 중심으로 하면 충분하고, 그에 더해 암호화 키나 매핑표 등 추가 정보의 관리에 관한 사항은 안전조치의무의 대상이 될 것이다. 반면 ② 프라이버시 보호의 내용과 기준을 설정함에 있어 신원정보를 넘어서는 부분까지 적극적인 보호가 필요하다는 견해를 취한다면 가명처리의 적정성 심사를 함에 있어 여러 가지의 위험을 입체적이고 종합적으로 고려하여 판단해야 한다. ①번 입장에 따르면 가명처리의 과정 자체는 상대적으로 수월할 수 있지만, 다양한 속성정보에 대한 통제는 쉽지 않을 수 있다. 따라서 그로 인한 재식별 위험을 방지하기 위한 절차적, 관리적 장치를 매우 세밀하게 마련할 필요가 있을 것이다. 반면 ②번 입장에 따르면 가명처리 적정성 심사는 가명정보 결합의 위험성을 망라한 종합적인 것이 된다.[41] 이 방식에 따르면 가명처리의 과정은 상당히 복잡한 과정이 될 수 있는 한편, 일단 적절하게 가명처리가 이루어진 데이터에 대해서는 추가적인 관리적 조치에 대한 부담이 줄어들 수 있다.

41 학계에서는, 프라이버시 공격의 의미를 적극적으로 고려할 필요가 있다는 의견이 계속적으로 제시되었다(Cynthia Dwork, Adam Smith, Thomas Steinke, and Jonathan Ullman, Exposed! A Survey of Attacks on Private Data, Annual Review of Statistics and Its Application Vol. 4:61-84 (March 2017)). 참고로, 공격은 다음의 방식으로서 구분하여서 설명할 수 있다. ① 공격의 주체는 내부자(insider adversary) 그리고 외부자(external adversary)이다. ② 공격의 목적은 정보주체의 신원정보의 식별(re-identification attack)이 일반적으로 논의되는데, 개인정보가 정보주체를 포함하는지 여부의 확인(membership attack)과, 민감한 속성의 추론(inference attack)도 주목할 필요가 있겠다. ③ 공격의 대상은 개인정보와 정보주체의 민감한 속성정보(sensitive attribute)가 우선적이고, 가명처리 비밀이 포함될 것이다. ④ 공격의 방법은, 예컨대, 무차별적 대입 공격(brute-force attack), 사전 공격(dictionary attack), 추측(guesswork)이 있는데, 식별정보와 가명정보의 크기와 매핑 함수의 설정이 그 효과를 결정할 것이다(ENISA, "Pseudonymisation Techniques and Best Practices", (2019)).

5) 민감정보의 가명처리

가명처리는 정보주체의 식별성과 관계된 속성을 대체하는 것이므로 속성정보의 민감성과 직접적인 관련이 있는 것은 아니다. 민감정보의 처리제한과 가명정보를 통한 처리제한에 관한 규정이, 가명처리한 민감정보의 이용과 제공을 금지한 것인지에 관하여 개정법의 조문에 기초하여 쉽게 결론은 내리기는 어렵다. 민감정보는 민감한 속성정보를 포함한 개인정보를 의미하는데, 가명정보는 개인정보이므로, 이는 민감한 속성정보를 포함한 가명정보의 처리에 제한이 있는지에 관한 문제이다(개정 개인정보 보호법 제23조 제1항, 제2조 제1호).[42] 이에 관해, 민감정보 그 중에서도 특히 건강에 관한 정보는 과학적 연구의 주요한 자료가 되는 경우가 빈번하므로 민감한 속성정보를 포함한 가명정보의 처리를 허용하는 것이 필요하다는 현실적인 입장이 있을 수 있다. 다른 한편, 민감정보는 그 민감성 때문에 특히 강력한 프라이버시 보호가 필요한 정보이므로, 가명처리 및 그 이후의 활용과정에서 프라이버시 침해 위험이 높아진다면 이는 허용될 수 없다는 주장도 제기될 수 있다. 현재로는 법률의 문언적 해석만으로는 답이 명료하지 않은 상황이다.

6) 다른 법과의 관계

가명처리와 관련하여 데이터 3법 이외의 법률에 규정된 내용과 어떻게 조화를 이를 것인지의 문제가 발생할 수 있다. 그 중 특히 생명윤리 및 안전에 관한 법률(이하 '생명윤리법')에 규정된 내용이 문제될 수 있다. 생명윤리법은 '익명화' 개념에 관해 규정하는 내용을 별도로 두고 있는데, 생명윤리법상의 익명화와 데이터 3법상의 '가명처리' 사이의 관계가 문제될 수 있다. 구체적으로, 생명윤리법은 익명화에 관해 '개인식별정보를 영구적으로 삭제하거나, 개인식별정보의 전부 또는 일부를 해당 기관의 고유식별기호로 대체하는 것'으로 규정하고 있다(생명윤리법 제2조 제19호). 이렇게 규정된 익명화 개념이 데이터 3법에 규정된 가명처리와 어떻게 같거나 다른지 명확하지 않다.

해석에 따라서는 이 조항은 개인정보 보호법과 충돌을 야기할 수도 있다. 더 넓게는, 생명윤리법 체계를 통해 진행되는 2차적 활용과 개인정보 보호법상의 제3자 제공에 관한 규율과의 관계에 대한 명확한 검토가 필요하다. 생명윤리법에서는 개인정보와 인체유래물의 제3자 제공에 대해 기관위원회의 심의를 요구하고 있는데, 데이터 3법이 이 과정에 변화를 요구하는지, 그 경우 어떤 변화가 필요할 것인지에 관해 추가적인 검토가 요구된다(생명윤리법 제18조, 제38조).

42 이동진, "개인정보 보호법 제18조 제2항 제4호, 비식별화, 비재산적 손해 - 이른바 약학정보원 사건을 계기로 -", 정보법학 제21권 제3호 (2017. 12.), 253-284면.

3. 가명정보의 보호

가. 총설

개인정보 보호법의 개정으로 인해 가명정보를 적극적으로 활용하고자 하는 개인정보처리자 및 이러한 활용을 촉진하고자 하는 정책당국은 개인정보 보호법의 개정취지를 달성함과 동시에 정보주체의 프라이버시 침해 위험을 최소화해야 하는 과제를 마주하게 되었다. 가명정보에 대하여 적절한 보호조치를 취하여 내·외부로부터의 공격에 적절히 대응할 수 있어야만 서로 상충하는 두 목표 사이에서 최적(optimal)의 선택을 할 수 있는 것이므로, 가명정보의 구체적인 보호 방법을 논하는 것은 개인정보 보호법의 개정취지에 반하는 것이 아닐 뿐 아니라 오히려 이에 부합하는 것이자 이를 달성하기 위하여 필수적인 작업이다. 개정 개인정보 보호법이 기존에 제29조[43]에서 개인정보처리자의 안전조치의무를 규정하고 있던 것에 더하여 제28조의4[44]에서 가명정보에 대한 안전조치의무를 추가로 규정한 것도 같은 맥락에서 이해될 수 있다.

따라서 이하에서는 가명정보에 대한 보호조치를 크게 기술적 보호조치와 관리적 보호조치의 두 가지로 나누어 살펴본다. 개정 개인정보 보호법 제28조의4 제1항은 "기술적·관리적 및 물리적 조치"라는 표현을 사용하고 있으나, 물리적 보호조치는 그 구체적인 내용에 따라 기술적 보호조치나 관리적 보호조치에 포함시켜 설명될 수 있기에 이를 별도의 목차로 분리하여 서술하지는 않기로 한다.[45]

가명정보에 대한 보호조치의 구체적 내용을 살펴보기에 앞서 한 가지 지적하고자 하는 것은, 가명정보를 처리하는 모든 개인정보처리자가 이하에서 제시하는 기술적·관리적 보호조치를 모두 동일하게 일괄적으로 이행해야 하는 것은 아닐 것이라는 점이다. 구체적인 규율의 내용을 정함에 있어서도, ① 가명정보를 처리하는 개인정보처리자의 유형 내지 규모, 그리고 ② 개인정보처리자가 실제로 처리하는 가명정보의 개별적인 특성(특히 유출 등 사고 발생시 정보주체의 프라이버시에 초래될 수 있는 침해의 중대성) 등을 종합적으로 고려하여 해당 개인정보처리자가 가명정보에 대하여 취해야 할 보호조치의 구체적 내용에 차등을 두어야 하고, 개인정보처리자 또한 개별 가명정보 처리 사안마다 위 요소들을 고려하여 의사결정을 할 필요가 있다.

43 개인정보 보호법 제29조(안전조치의무) 개인정보처리자는 개인정보가 분실·도난·유출·위조·변조 또는 훼손되지 아니하도록 내부 관리계획 수립, 접속기록 보관 등 대통령령으로 정하는 바에 따라 안전성 확보에 필요한 기술적·관리적 및 물리적 조치를 하여야 한다.

44 개정 개인정보 보호법 제28조의4(가명정보에 대한 안전조치의무 등) ① 개인정보처리자는 가명정보를 처리하는 경우에는 원래의 상태로 복원하기 위한 추가 정보를 별도로 분리하여 보관 · 관리하는 등 해당 정보가 분실·도난·유출·위조·변조 또는 훼손되지 않도록 대통령령으로 정하는 바에 따라 안전성 확보에 필요한 기술적·관리적 및 물리적 조치를 하여야 한다. ② 개인정보처리자는 가명정보를 처리하고자 하는 경우에는 가명정보의 처리 목적, 제3자 제공 시 제공받는 자 등 가명정보의 처리 내용을 관리하기 위하여 대통령령으로 정하는 사항에 대한 관련 기록을 작성하여 보관하여야 한다.

45 참고로 GDPR은 "기술적 및 관리적 조치(technical and organisational measure; TOM)"라는 표현을 사용하고 있다. GDPR Recital (29) 등 참조.

나. 기술적 보호조치

가명정보 또한 개인정보이므로 안전조치의무의 일환으로 규정되어 있는 기술적 보호조치가 적용된다(개정 개인정보 보호법 제28조의4, 제29조). 이에 더하여 가명처리 과정에서 통상적인 개인정보의 처리에 있어서는 고려되지 않을 추가적인 기술적 보호조치가 필요할 수 있다. 이러한 가명정보 및 가명처리의 법적 쟁점들을 잘 이해하고 원활하게 해결하기 위해서는 무엇보다도 구체적인 가명처리 기법과 개념에 대한 이해가 선행되어야 한다.[46] 이하에서는 가명정보 및 가명처리의 영역에 있어 법과 기술 간의 연결고리를 마련하는 작업의 일환으로서 가명처리의 기본 구조와 가명처리 기법을 풀어 설명한다.

1) 가명처리의 기본 구조

위에서 본 것과 같이 개정 개인정보 보호법상 가명처리는 '개인정보의 일부를 삭제하거나 일부 또는 전부를 대체하는 등의 방법으로 추가 정보가 없이는 특정 개인을 알아 볼 수 없도록 처리하는 것'으로 정의된다(동법 제2조 제1의2호). 이는 결국 이름, 이메일 주소, 휴대전화번호 등 특정 정보주체를 식별할 수 있는 정보(이를 식별자(identifier)라 한다)에 대하여 그 중 일부를 삭제하거나 일부 또는 전부를 대체하는 등의 방법으로 만들어진 가명(pseudonym)을 대응시키는 것을 기본적으로 의미한다. 가령, 자연인 김갑동의 식별자인 이메일 주소 kkd@snu.ac.kr에 @을 기준으로 하여 앞부분(local)에서는 각 문자를 알파벳 순서상 그로부터 2번째 뒤의 문자로 대체 하고 뒷부분(domain)에서는 첫 문자를 제외하고 나머지를 삭제하여 만든 mmf@s라는 가명을 대응시키는 것이 가명처리의 한 사례이다. 이러한 개념정의상 가명처리는 그 자체로 하나의 함수(function) 내지 매핑(mapping) 과정으로 이해될 수 있는데, 이때 이용되는 함수를 가명처리 함수(pseudonymization function)라 한다.

모든 가명처리 함수가 완전히 '복원 가능(recoverable)'해야 하는 것은

46 이를 위해서는 암호학의 기초적인 개념을 이해하는 것이 도움된다. 후술하는 바와 같이 가명처리 기법이 꼭 수학적인 의미에서의 암호화(encryption)로 한정되는 것은 아니지만, 그 내용은 대체로 암호학의 기초적인 개념들(가령, 암호화와 복호화(decryption) 알고리즘과 이에 사용되는 키(key) 등)의 활용 내지 응용으로 설명될 수 있기 때문이다. 또한, 흔히 언급되는 가명정보에 대한 재식별 공격(re-identification attack) 시나리오들은 암호학에서 상정하고 있는 각종 공격 유형들(가령, 후술하는 무차별 대입 공격(brute force attack), 사전 대입 공격(dictionary attack) 등)의 일부이다.

아니다(개정 개인정보 보호법 제2조 제1호 다목 참조). 그러나 실무적으로는 개별적 필요에 따라 각 가명으로부터 원래의 식별자를 복원해낼 수 있도록 하는 경우가 종종 나타나는데, 이러한 경우 개별 식별자를 복원해낼 수 있도록 하는 가명처리 함수 관련 정보 일체를 통틀어 '가명처리 비밀(pseudonymization secret)'이라 한다.[47] 가명처리 비밀은 다양한 형태를 띨 수 있는데, 위 사례의 경우에 이메일 주소 앞부분의 문자 대체 간격이 하나의 예가 될 수 있고, 각 식별자와 그에 대응되는 가명을 표로 나타낸 가명처리 매핑표(pseudonymization mapping table)가 있다면 그러한 표가 또 다른 예가 될 수 있다.

가명처리 비밀과 개정 개인정보 보호법상 '원래의 상태로 복원하기 위한 추가 정보'는 어떠한 관계에 있는지 해석상 논란이 있을 수 있다(동법 제2조 제1호 다목 및 제1의2호, 제28조의4 제1항). 가명처리 비밀 이외의 정보 중에도 가명과 결합하여 원래의 식별자를 복원하는 데 사용될 수 있는 정보, 즉, 보조정보가 있을 수 있기 때문이다. 그런데 가명처리 비밀과 달리 보조정보는 일반적으로 개인정보처리자의 통제 영역을 벗어난 것이므로 이러한 외부 정보까지 '원래의 상태로 복원하기 위한 추가 정보'에 포함된다고 본다면, 가명정보를 처리하는 경우 개인정보처리자가 위와 같은 추가 정보를 별도로 분리 하여 보관·관리해야 함을 규정한 개정 개인정보 보호법 제28조의4 제1항은 어떻게 준수할 수 있을지에 관한 또 다른 어려움이 나타난다. 다른 한편, 가명처리 비밀만을 고려하는 것으로 법적 요건이 충족되는 것으로 해석한다면, 그 이외의 방식과 경로를 통한 재식별 가능성은 실질적으로 무시되는 결과가 나타날 수 있어서 이를 통해 프라이버시 보호에 한계가 나타날 가능성도 있다.

개념적으로는, 가명처리 함수와 관련된 것으로서 개인정보처리자가 기술적·절차적으로 통제할 수 있는'가명처리 비밀'그리고 그러한 통제가 어려운 것으로서 특정 식별자의 가명처리에 있어 맥락(context) 내지 환경(situation)에 따라 달리 나타날 수 있는 요소인 보조정보를 구분하여, 이러한 다양한 요소를 적절하게 고려하는 것이 이상적인 것이라 할 수 있다. 다만 이를 현실적으로 어떻게 구현할 것인지에 관하여 법해석에서의 어려움 및 실무적인 어려움이 있을 수 있다. 이하에서는 두 개념의 구분이 필요한 경우에, 개정 개인정보 보호법상의 '원래의 상태로 복원하기 위한 추가 정보' 내지 '추가 정보'라는 표현보다는 '가명처리 비밀'과 '외부 정보'라는 용어를 사용한다.

47 연구문헌에서는 '가명처리 비밀(pseudonymization secret)'이라는 용어보다는 '가명처리 키(pseudonymization key)'라는 용어가 좀 더 일반적으로 사용되는 것으로 보인다. 그러나 가명처리 함수에서 식별자 외의 입력값으로서 키(key)가 요구되는 방식이 항상 이용되는 것은 아니고, 나아가 후술하는 바와 같이 식별자의 복원은 가명처리 매핑표(pseudonymization mapping table) 등에 의하여도 이루어질 수 있기 때문에 '가명처리 키'보다는 '가명처리 비밀'이라는 좀 더 포괄적인 용어를 사용하기로 한다.

가명처리 비밀은 가명에 대응하는 원래의 식별자를 복원해냄에 있어 핵심적인 역할을 수행하므로 개인정보처리자는 이를 암호화하여 별도의 정보저장매체에 저장하는 등의 기술적·물리적 조치를 취해야 할 뿐만 아니라, '신뢰할 수 있는 제3자(Trusted Third Party)'에게 그 관리의 전부 또는 일부를 위탁하거나 내부 통제를 강화하는 등 적절한 관리적 조치도 취해야 한다. 개정 개인정보 보호법 제28조의4 제1항은 이를 명문으로 규정하고 있는 한편, 그 구체적인 내용은 대통령령에 위임되어 있다.

2) 가명처리 기법

가) 가명처리 기법의 분류 - 구조(structure) 보존의 정도

많은 경우 식별자들은 서로 일정한 관계를 맺고 있는 것으로 파악할 수 있다. 가령 생년월일은 서로 선후 순서를 비교할 수 있고, 이름은 사전식 순서(lexicographic order)에 따라 정렬될 수 있다. 나아가 많은 경우 식별자들은 자체적으로 일정한 형식으로 구성되어 있다. 가령 이메일 주소는 @을 기준으로 앞부분과 뒷부분으로 나뉘며, 주민등록번호는 일정한 규칙에 따라 부여된다. 또한 문자나 숫자의 개수가 사전에 정해져 있는 경우도 많다.

이러한 식별자들 간의 관계나 식별자들의 개별 형식 등을 추상적으로 하나의 구조(structure)라고 할 수 있다. 그런데 가명처리는 그 과정에서 일정한 정보손실이 초래되는 경우가 많고, 데이터의 구조 자체에 변화가 나타나기도 한다. 또한, 이러한 변화가 정보에 왜곡을 가져올 수도 있다. 가령, 이메일 주소의 사례를 생각해 보자. 데이터베이스에 kkd@snu.ac.kr과 kkd@seoul.go.kr이 포함되어 있다고 할 때 식별자의 차원에서 전자는 후자보다 사전식 순서상 뒤에 있다. 그러나 앞서 정의한 방식으로 가명처리를 할 경우 두 개의 이메일 주소에 대해 그 결과는 모두 mmf@s로 같아서 더 이상은 이 둘의 사전적 순서를 비교할 수 없게 된다(이를 '가명 간 충돌(collision)'이 발생한 경우라 한다). 따라서 만약 위 사례에서 가명정보를 서로 사전식 순서에 따라 비교하는 것이 목적이었다면, 이 방식의 가명처리 기법을 활용해서는 곤란하다. 이러한 경우에는, 예를 들어, @ 뒷부분을 삭제하지 않고 남겨두거나 @ 앞부분을 처리한 것과 마찬가지로 각 문자를 알파벳 순서상 그로부터 일정한 간격 뒤의 문자로 대체하는 등의 방식을 이용하여 식별자 간에 존재하는 사전식 순서 구조를 보존하는 가명처리 기법을 활용해야 할 것이다.

이처럼 가명처리 기법마다 처음의 데이터를 보존하는 구조 내지 그 보존의 정도는 다를 수 있는바, 가명처리 기법은 이에 따라 다양한 방식으로 분류될 수 있다. 그 중 어떤 가명처리 기법을 선택하여 어떤 방식으로 구현할 것인지는 결국 가명처리의 목적을 고려하여 판단되어야 할 문제이다. 일반적으로는 더 상세한 구조를 높은 수준으로 보존할수록 가명정보를 더욱 유용하게 활용할 수 있는 가능성이 높아지겠지만, 다른 한편 잠재적 공격자를 고려할 때 원래의 식별자에 관한 정보를 더 많이 노출하는 것은 개인정보 보호를 달성하는 데에 어려움을

야기할 수 있다. 구조 보존의 정도에 따라 가명처리 기법을 분류하는 것은 결국 가명처리에 따른 가명정보의 유용성(utility)과 개인정보 보호(data protection) 사이의 상충관계(trade-off)를 보여주는 것이다.[48]

나) 가명처리 기법들[49]

(1) 암호학적 해시함수(cryptographic hash function)

일반적으로, 암호학적 해시함수(이하 "해시함수")는 임의의 길이의 메시지(message)를 입력받아 고정된 길이의 메시지(해시값(hash value) 또는 메시지 요약(message digest)이라 한다)를 산출해내는 함수로서, 개인정보의 가명처리를 위해서도 활용가능성이 높을 것으로 예상된다. 보편적으로 쓰이는 해시함수들의 계산 알고리즘은 일반 공개가 되어 있으므로, 누구나 이를 활용할 수 있다. 현재 널리 쓰이는 알고리즘의 예로 SHA-256를 들 수 있는데, 이 방식은 해시값의 계산이 비교적 빨라 공인인증서, 블록체인 등에서 널리 활용되고 있다.

다른 가명처리 기법과의 결합 없이 해시함수만을 이용하여 가명처리를 함에 있어 서는 크게 두 가지 방식을 생각할 수 있다. 하나는 식별자(id)만을 입력하여 그 해시값 ($h(id)$)을 가명으로 하는 것이고, 다른 하나는 식별자뿐만 아니라 별도로 정한 키(k) 또는 무작위로 생성한 솔트($salt$)까지 입력받아 그 해시값(가령, $h(id||k)$)나 $h(id||salt)$) 을 가명으로 하는 것이다.

해시함수의 계산 알고리즘은 모두 공개되어 있으므로, 이러한 두 가지 가명처리의 방법 중에서 첫 번째 방법은 두 번째 방법에 비해 잠재적 공격자에 의한 무차별 대입 공격(brute force attack),[50] 사전 대입 공격(dictionary attack)[51] 및 통계적 공격(statistical attack)[52] 등에 취약할 수밖에 없다. 따라서 해시함수를 이용하여 가명 처리를 하는 경우에는 식별자를 특정 해시함수에 대입하여 계산하는 첫번째 방식보다는, 솔트 등을 포함하는 두 번째 방식을 이용하는 것이 재식별의 리스크를 낮추는 데에 더욱 도움이 될 것이다.

(2) 암호화(encryption)

개념적으로 암호화는 일정한 메시지('평문')를 일정한 사람들만이 그 의미를 파악할 수 있도록 다른 형태의 메시지('암호문')로 변환하는 것을 의미한다. 실제로 데이터를 활용하는 경우에, 변환은 암호화 함수(encryption function)에 평문과 암호화 키(encryption key)를 대입하는 방식으로 이루어지는 것이 보통이다. 한편, 암호문을 다시 평문으로 변환하는 과정은 복호화(decryption)라 하는데, 이 또한 대부분의 경우 복호화 함수(decryption function)에 암호문과 복호화 키(decryption key)를 대입하는 방식으로 이루어진다.

이러한 암호화는 해시함수를 이용한 가명처리와 그 구조가 유사하다. 실제로 암호학에서 개발되어 있는 다양한 암호화 기법들을 활용하여 식별자를 암호화하는 방식으로 가명처리를 수행할 수 있다. 암호체계는 크게 대칭키 암호체계(symmetric key cryptosystem)와 공개키 암호체계(public key cryptosystem)의 두 가지로 나뉘는데, 가명처리에 있어서는 일반적으로 공개키 암호체계보다는 대칭키 암호체계를 활용하는 것이 더 효과적일 것으로 평가된다. 별도의 조치를 취함이 없이 단순히 식별자를 공개키로 암호화하여 가명으로 활용하는 경우, 암호화 함수와 공개키가 모두 공개되어 있으므로 잠재적 공격자의 무차별 대입 공격, 사전 대입 공격 및 통계적 공격 등에 취약할 것이기 때문이다. 또한 일반적으로 공개키 암호체계는 대칭키 암호체계에 비해 그 암호화 및 복호화에 있어 더 많은 연산량을 요구하므로, 정보량이 많거나 체계가 복잡한 데이터베이스를 가명처리하기에는 적합하지 않은 면이 있다.

암호화를 활용한 가명처리 기법은 가명처리 매핑표 전부를 저장할 필요 없이 키와 가명들만을 저장해두면 이를 복호화 함수에 대입하여 원래의 식별자를 복원해낼 수 있으므로, 정보량이 많거나 체계가 복잡한 데이터베이스를 가명처리하는 데에 유용하게 쓰일 수 있다. 그러나 이 또한 암호체계에 대한 각종 공격으로부터 자유롭지 않으므로, 실제로 가명처리를 수행함에 있어서는 개별 암호화 기법의 암호학적 특징과 취약성을 고려하여 판단하여야 한다.

다. 관리적 보호조치

정교한 암호화 알고리즘을 활용하는 등 적절한 기술적 보호조치를 취하는 것을 통해 가명정보를 충분히 보호할 수 있다고 생각하는 경우도 있을 수 있다. 그러나 이는 비현실적인 바람일 뿐 아니라, 오히려 가명정보의 보호를 위한 또 다른 축으로서 관리적 보호조치가 갖는 중요성을 간과함으로써 그 보호에 있어 중대한 위험을 초래할 우려가 있다. 아무리 강력한 기술적 보호조치를 취한다

48 이동진, "개인정보 보호법 제18조 제2항 제4호, 비식별화, 비재산적 손해 - 이른바 약학정보원 사건을 계기로 -", 정보법학 제21권 제3호 (2017. 12.), 253-284면.

49 이하의 내용은 ENISA, "Recommendations on Shaping Technology according to GDPR Provisions" (2018. 11.), 19-30면 등을 보완·정리한 것이다.

50 무차별 대입 공격이란, 수많은 값들을 무차별로 대입해가며 공격을 수행하는 것을 뜻한다. 가령, 어떠한 가명이 식별자를 SHA-256에 대입하여 생성되었다는 사실을 알고 있는 공격자가 무작위로 문자열을 선택한 뒤 이를 위 해시함수에 대입해가며 해당 가명에 대응하는 원래의 식별자를 재식별하기 위해 시도하는 것이 그 예이다.

51 사전 대입 공격이란, 사전에 등장하는 단어들과 같이 전형적으로 쓰이는 값들을 대입해가며 공격을 수행하는 것을 뜻한다. 이는 무차별 대입 공격을 개선시킨 것으로서, 특히 보조정보와 결합되었을 때 더 큰 효과를 발휘한다.

52 통계적 공격이란, 맥락에 따라 다소 다른 의미로 쓰이는 경우가 있기는 하나, 일반적으로는 평문과 암호문의 통계적 구조를 비교 분석함으로써 공격을 수행하는 것을 뜻한다. 셜록 홈즈 시리즈의 유명한 에피소드 "춤추는 사람들(The Adventure of the Dancing Men)"에 등장하는 빈도 분석(frequency analysis)이 그 대표적인 사례이다. 가령 성(姓)의 해시값을 그 가명으로 한 경우, 가명 데이터베이스가 충분히 크다는 전제 하에 공격자는 위 데이터베이스에서 가장 많이 등장하는 가명은 높은 확률로 김(金)에 대응하는 것이라고 추론해낼 수 있을 것이다.

하더라도 결국 키나 가명처리 매핑표와 같은 가명처리 비밀을 생성·관리하는 것은 사람이므로, 가명처리 등 가명정보 처리의 전 과정에 있어 개인정보 보호책임자를 지정하고(개인정보 보호법 제31조), 업무를 분장하며, 주기적으로 모니터링을 실시하는 등 적절한 관리적 보호조치를 취해야 하고, 이를 위한 절차를 미리 마련해야 한다. 아무리 정교한 암호화 알고리즘을 사용하여 가명처리를 수행하더라도, 관리적 보호조치가 취약하여 그 키가 접근권한의 범위를 넘어 노출되거나 하면 그러한 가명처리는 유명무실한 것일 수밖에 없다.

개정 개인정보 보호법은 가명정보를 처리하는 개인정보처리자가 취해야 할 관리적 보호조치의 세부 내용을 대통령령에 위임하고 있고, 동법 시행령 일부개정령안 제29조의5가 이를 어느 정도 구체화하고는 있으나, 향후 하위 법령 및 가이드라인 등 추가적인 방식을 통해 더 많은 논의와 구체화가 필요한 상황이다. 이러한 논의에 있어 시사점을 제시하기 위해, ① 경영학, 회계학 등에서 다각도로 연구되어 온 내부통제(internal control) 및 외부감사(external audit) 관련 제도를 응용하여 맥락에 맞게 변용·수용할 수 있는 가능성을 제시해 본다. 그리고 ② 신뢰할 수 있는 제3자(TTP)를 활용하는 방안을 소개하며 적용가능성을 검토해 본다.

1) 내부통제와 외부감사

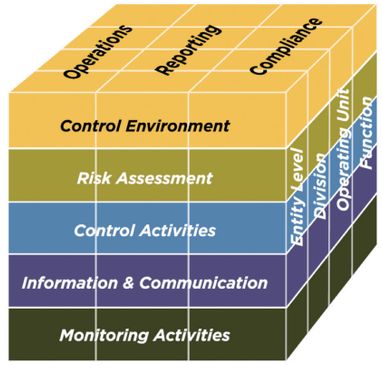
가) 내부통제

회계감사의 맥락에서 내부통제는 "재무보고의 신뢰성, 경영의 효과성 및 효율성, 그리고 관련 법규의 준수에 관련된 기업의 목적 달성에 관한 합리적 확신을 제공할 목적으로 지배기구, 경영진 및 기타의 인원에 의해 설계, 실행, 유지되고 있는 절차"[53]로 정의된다. 이 개념을 가명정보의 맥락에 어렵지 않게 응용할 수 있다. 특히 조직 규모가 크고 운영 과정이 복잡해질수록, 관련 법규를 준수하면서 운영의 효과성 및 효율성을 달성하기 위해서 적절한 내부통제 시스템을 마련하는 것은 필수적이다.

개정 개인정보 보호법에 따라 가명정보를 처리하고자 하는 개인정보처리자 또한 마찬가지이다. 개인도 개인정보처리자가 될 수 있지만, 가명처리에 소요되는 자원 규모 등에 비추어보았을 때 현실적으로 가명정보의 처리 및 그에 대한 통제가 문제되는 것은 주로 개인정보처리자가 공공기관, 법인 또는 단체인 경우일 것이므로 적절한 내부통제 시스템을 구축하고 운용할 필요가 있다. 개인정보 보호법에 따라 개인정보처리자는 개인정보 보호책임자를 지정하여야 하는데, 개인정보 보호책임자의 업무 중 하나로 '개인정보 유출 및 오용·남용 방지를 위한 내부통제 시스템의 구축'이 포함되어 있는 것도 같은 맥락에서 이해될 수 있다(개인정보 보호법 제31조 제1항, 제2항 제4호).

53 회계감사기준(2018년 개정) 감사기준서 315 문단 4(c); International Standard on Auditing(ISA) 315(Revised 2019) 문단 12(m).

다만 가명정보 관련 내부통제 시스템을 구축하여 운용함에 있어서는 단순히 개인정보의 일반적 특성들을 고려하는 것으로는 충분하지 않고, 가명처리 및 가명정보의 기본 개념, 해당 개인정보처리자가 실제로 처리하는 가명정보의 특성과 그에 따른 재식별 내지 프라이버시 침해 위험 등을 정확히 파악하여 반영해야 한다. 내부통제의 구성요소로는 통상 ① 통제환경(control environment), ② 조직의 위험평가절차(risk assessment process), ③ 정보시스템 및 커뮤니케이션(information system and communication), ④ 통제활동(control activities), ⑤ 통제의 모니터링(monitoring)의 다섯 가지를 꼽을 수 있는데,[54] 이하에서는 가명정보 관련 내부통제의 구체적 내용을 위 각 구성요소별로 나누어 살펴보기로 한다.



[그림 1] COSO 내부통제 모델[55]

우선 ① 통제환경은 "내부통제와 그 중요성에 관한 지배기구 및 경영진의 태도와 의식 및 행동"[56]을 비롯하여 내부통제의 전반적인 환경을 의미한다. 이는 내부통제의 기초가 되는 것으로서, 현실적으로 가장 중요한 요소 중 하나이다. 경영진이 가명처리 및 그와 관련된 통제장치를 단지 실무진 사이의 기술적인 문제로 치부할 경우, 효과적인 내부통제는 사실상 불가능하게 되기 때문이다. 특히 ⓐ 가명정보의 유용성과 개인정보 보호 간에는 일정한 상충 관계가 존재하는 것이 일반적인데, 그 상충관계의 내용을 구체화하고 이로부터 개별 조직이 추구하는 유용성이 어떤 것인지 새로이 검토하는 등의 지속적인 노력이 필요하다는 것과, ⓑ 한 차례 가명처리를 수행하였다고 해서 그것만으로

54 회계감사기준 감사기준서 315 문단 14-24; ISA 315 문단 12(m); Committee of Sponsoring Organizations of the Treadway Commission(COSO), Internal Control - Integrated Framework, Executive Summary (2013. 3.) 등 참조.

55 COSO(위 각주), 6면.

56 회계감사기준 감사기준서 315 문단 A76; ISA 315 부록 3 문단 4-6.

가명정보의 보호가 달성될 수 있는 것이 아니고, 가명정보 처리의 전 과정에 있어 내부통제 시스템을 구축하여 지속적으로 운용하는 등 각종 관리적 보호조치를 취해야 한다는 점을 명확히 하는 것이 중요하다.

② 조직의 위험평가절차는, ⓐ 관련 위험을 식별하고, ⓑ 해당 위험의 발생가능성을 평가하며, ⓒ 그에 대처하기 위한 행동을 결정하기 위한 조직 내부의 절차를 의미한다.[57] 이는 가명정보의 맥락에서 특히 중요한데, 개인정보처리자는 자신이 보유한 가명정보에 대한 다양한 재식별 공격 시나리오들을 선제적으로 상정하여 그에 따른 위험을 파악 및 평가하고 그에 대한 대응책을 결정하기 위한 내부 절차를 수립하여야 한다. 이 과정에서 가명정보의 재식별 위험을 보다 정확하게 평가하기 위하여 모의 재식별 공격을 수행하는 등으로 가명정보의 보안성을 주기적으로 시험하고, 취약점이 발견된 경우 추가적인 보안 조치를 취할 수 있는 전문가, 즉 화이트 해커(white hacker)들로 구성된 팀을 별도로 꾸려 운영하는 것이 효과적일 수 있다.

③ 정보시스템 및 커뮤니케이션은 조직 내부에서 관련 정보가 생성되고 유통되어 일부는 파기되고 다른 일부는 내부적으로 기록되며 그 중에서도 선별된 일부는 외부에 공개되는 등 정보 관리의 전반적인 체계를 의미하는데, 특히 ⓐ 조직에서 유의미한 활동(가령 유통을 주 사업으로 영위하는 기업의 경우 재고자산의 매입거래 등)이 개시, 기록, 처리, 보고되는 절차, ⓑ 관련 기록과 증빙서류의 보관 방법, ⓒ 경영진과 지배기구 간 커뮤니케이션 및 ⓓ 규제기관 등 외부와의 커뮤니케이션 등이 그 주요 내용이다.[58] 이는 가명정보에 있어서도 마찬가지로 중요한데, 가명정보를 처리하는 개인정보처리자는 특히 그 처리 목적, 처리 및 보유기간, 추가정보의 이용 및 파기 등 관련 중요 정보를 기록하여 문서화하고, 해당 기록을 보관해야 한다(개정 개인정보 보호법 제28조의4 제2항, 동법 시행령 일부개정령안 제29조의5 제3항). 또한 가명정보의 처리 목적이 달성되거나 가명정보 보유 기간이 경과한 때에는 가명정보를 지체 없이 파기해야 한다(위 시행령 일부개정령안 제29조의5 제3항). 그리고 이를 위하여, 개개 개인정보취급자에 대한 로그 시스템 및 사후 검증 절차 등을 상세하게 마련해 두는 것이 현실적으로 중요할 수 있다.

④ 통제활동은 "경영진의 지시가 확실하게 이행되도록 도와주는 정책과 절차"로서 승인(authorization), 성과검토(performance review), 정보처리(data processing), 물리적 통제(physical control), 업무분장(segregation of duties) 등을 포함한다.[59] 통제활동은 조직의 활동에 가장 직접적인 영향을 미치는 것으로서 내부통제의 핵심 구성요소에 해당한다. 따라서 가명정보의 맥락에서 각 통제활동이 어떻게 설계·수행되어야 하는지를 살펴본다.

우선 승인이란 조직에서 일정한 활동이 이루어지는 경우 "적절한 책임을 가진 자가 해당 활동에 대해서 사전 또는 사후적 인·허가를 하는 절차"를 말하는데,[60] 가명정보에 있어서도 마찬가지로 적절한 승인절차를 설계하여 운영할 필요가 있다. 가명정보의 맥락에서 승인의 대상이 되는 것으로는 가명처리의 대상, 구체적인 가명처리 알고리즘, 가명정보의 처리 방법 등 다양한 사항들을 상정할 수 있다. 그런데 개별 사안에서 그 승인 여부를 결정함에 있어서는 법적 지식뿐만 아니라 가명처리 기법 및 관련 기술들에 대한 심도 있는 이해가 필요하므로 개인정보 보호책임자 개인이 전적으로 이를 결정하도록 하는 것은 적절하지 않을 수도 있다. 그 경우 의료 분야에서 '인간대상연구'등에 관하여 연구의 윤리적·과학적 타당성 등을 심의하기 위해 각 기관별로 독립적으로 설치하는 기관생명윤리위원회 또는 임상시험심사위원회(institutional review board; IRB)[61] 등을 참조하여, 개인정보처리자 내부적으로 데이터심의위원회(data review board; DRB)를 설치하는 것을 고려할 수 있다. 데이터심의위원회를 설치하는 경우 그 기본적인 역할은 가명정보 처리의 전 과정에서 발생할 수 있는 문제들을 사전·사후적으로 파악하고 이를 적절히 승인하거나 불허하는 것일 테지만, 개별적 필요에 따라 위 위원회로 하여금 가명정보에 관하여 조직이 전반적으로 공유해야 할 정책적 방향이나 기준을 정하는 역할을 수행하도록 할 수도 있다. 또한 가명처리의 과정에 요구되는 통제환경을 설계·구축하는 역할을 수행할 수도 있다.

다음으로 가명정보의 맥락에서 중요한 통제활동으로는 물리적 통제를 꼽을 수 있는데, 이는 통상 "자산과 기록의 접근에 대한 안전설비와 같은 적절한 안전장치 등 자산의 물리적 보안"이나 "컴퓨터 프로그램과 데이터 파일에 대한 접근의 승인"을 포함한다.[62] 물리적 통제는 통제활동 중에서도 가장 직접적인 통제 내지 제약으로서, 적절히 이루어질 경우 큰 효과를 발휘할 수 있다. 따라서 가명정보를 처리하는 개인정보처리자는 ⓐ 가명처리 비밀을 출력하여 물리적 형태로 보관하는 경우에는 이를 잠금장치가 있는 안전한 장소에 보관함과 동시에 해당 장소에 대한 출입통제 절차를 수립·운영하여야 할 것이고, ⓑ 가명처리 비밀을 USB 등 보조저장매체에 저장하여 전자적 형태로 보관하는 경우에는 이를 위와 같이 안전한 장소에 보관함과 동시에 해당 매체의 복제나 반출·입 통제를 위한 보안대책을 마련하여야 할 것이다(개인정보의 안전성 확보조치 기준 제11조 참조). 군(軍)이나 정보기관 등에서 여러 등급을 나누어 비밀취급인가(security

57 회계감사기준 감사기준서 315 문단 15; ISA 315 문단 22 참조.

58 회계감사기준 감사기준서 315 문단 18-19; ISA 315 부록 3 문단 15-19 참조.

59 회계감사기준 감사기준서 315 문단 A96; ISA 315 부록 3 문단 20-21 참조.

60 이창우·송혁준·전규안·권오상, 회계감사 Study Guide (제6판), 경문사 (2019), 제8장 8면 참조.

61 생명윤리법 제10조, 의약품 등의 안전에 관한 규칙 [별표 4] 의약품 임상시험 관리기준 제2호 오목 참조.

62 회계감사기준 감사기준서 315 문단 9; ISA 315 부록 3 문단 20 참조.

clearance)를 부여하는 것을 참조하여, 가명정보에 있어서도 필요하다면 단계별로 접근권한을 달리 부여할 수 있다. 나아가 정보주체의 프라이버시 침해 우려가 특히 높은 정보를 가명처리하여 보유하고 있는 경우에는 특별한 사정이 없는 한 해당 가명정보가 반출되어 다른 저장매체에 저장될 수 없도록 허가받은 단말기(terminal)를 통해서만 이를 열람할 수 있도록 하는 것도 가능하다. 또는 가상환경(virtual environment)에서만 데이터에 대한 처리가 가능하도록 시스템을 구축할 수도 있다.

한편, 통상 "거래의 승인, 거래의 기록, 자산의 보호에 대한 책임을 서로 다른 자에게 배정하는 것"이라고 정의되는 업무분장 또한 가명정보의 맥락에서 중요한 통제활동에 해당하는데, 이는 "누구든지 일상적인 직무수행 중 오류나 부정을 범하고 동시에 이를 은폐시킬 수 있는 위치에 있을 수 있는 기회를 감소시키기 위한 것"이다.[63] 따라서 가명정보를 처리하는 개인정보처리자는 가능한 한 ⓐ 가명정보의 개별 처리 목적을 설정하고 그에 따른 처리 여부를 승인하는 자, ⓑ 구체적으로 가명처리를 수행하는 자, ⓒ 가명처리를 비롯하여 가명정보의 처리 과정을 기록하는 자, ⓓ 가명처리에 따라 생성된 가명정보를 실제로 활용·처리하는 자를 각기 달리 정하여야 한다. 아래에서 설명하는 '신뢰할 수 있는 제3자(TTP)의 활용' 또한 기본적으로는 이러한 업무분장의 사고에서 비롯된 것이다.

마지막으로 ⑤ 통제의 모니터링은 "지속적으로 내부통제의 성과에 대한 효과성을 평가하는 절차"로, "적시에 통제의 효과성을 평가하고 필요한 개선조치를 취하는 것을 포함"한다.[64] 따라서 가명정보를 처리하는 개인정보처리자는 처리의 전 과정에 걸쳐 가명정보가 효과적으로 보호되고 있는지 여부를 지속적으로 평가해야 한다. 특히 관련 기술이 지속적으로 발전함에 따라 한때는 안전하다고 여겨졌던 가명정보가 추가 정보 없이도 재식별될 수 있는 정보로서 더 이상 가명정보에 해당하지 않게 될 위험이 상존하므로, 개인정보처리자는 이러한 위험을 항상 인식하고 선제적으로 대응하여야 한다. 이 과정에서 필요하다면 최신 보안 및 공격 기술에 익숙한 화이트 해커들로 하여금 주기적으로 모의 재식별 공격을 수행하는 등의 방식으로 가명정보의 보안성을 주기적으로 확인하는 과정이 필요할 수 있다.

나) 외부감사

내부통제를 아무리 효과적으로 설계·실행·유지한다 하더라도, 이는 그 특성상 일정한 한계를 가질 수밖에 없다. 내부통제의 한계로는 통상 ① 의사결정에 있어 인적인 판단이 잘못될 수 있는 가능성 및 인적 오류, ② 2명 이상의 공모

63 위 각주 참조.

64 회계감사기준 감사기준서 315 문단 A106; ISA 315 부록 3 문단 10-14 참조.

가능성, ③ 경영진이 부적합하게 내부통제를 무시할 가능성 등이 지적되는데,[65] 외부감사는 이러한 한계를 보완하기 위해 필요한 것이다. 가명정보의 맥락에서 외부감사란 '가명정보의 생성, 저장, 보유, 가공, 편집 등 처리의 전 절차에 걸쳐 충분한 보호조치가 이루어지는 등 개인정보 보호법의 관련 규정들이 준수되었는지 여부를 개인정보처리자 외부의 독립된 감사인이 검증하는 절차'를 의미한다고 볼 수 있다. 다만 개정 개인정보 보호법은 가명처리와 관련하여 외부감사에 관한 별도의 규정을 두고 있지는 않다.

만약 가명정보를 처리하는 개인정보처리자에 대한 외부감사제도를 고려한다면, 가명정보에 대한 감사기준을 마련하고 감사인의 요건이나 자격기준을 정하는 한편 이해충돌을 방지할 수 있는 각종 장치들을 마련해야 할 것이다. 한편 외부감사를 두는 것은 내부통제 시스템을 구축·운용하는 것에 비해 더욱 높은 비용을 수반할 것을 고려하여, 제도를 더욱 세심하게 설계할 수 있을 것이다.[66] 즉, ① 가명정보를 처리하는 개인정보처리자의 유형 내지 규모 및 ② 개인정보처리자가 실제로 처리하는 가명정보의 개별적인 특성 등 다양한 요소들을 고려하여 규율의 체계 및 내용을 달리 정할 수 있을 것이다. 이와 달리, 개인정보 보호법에 이미 마련되어 있는 개인정보 보호 인증제도를 응용 또는 확대하는 것을 또 다른 대안으로 고려해 볼 수도 있다(개인정보 보호법 제32조의2). 이 방식은 외부감사제도에 비해 제약이 상대적으로 적은 수단으로서 가명정보를 처리하는 개인정보처리자가 자발적으로 적절한 보호조치를 취하도록 유도하는 유용한 수단으로 활용될 수도 있을 것이다.

2) 신뢰할 수 있는 제3자(TTP)의 활용

가) 신뢰할 수 있는 제3자(TTP)의 의미

개정 개인정보보호법은 가명처리에 대한 개념정의를 하고 있지만 가명처리의 방법을 구체적으로 제시하고 있지는 않다. 또한 가명정보의 사후적 관리 의무에 대한 규정도 개인정보에 대한 일반적인 사후적 보호 의무 및 추가 정보의 분리보관 의무 이외에는 더 상세한 내용을 담고 있지는 않다. 그런데 실제로 개인정보의 가명처리를 하는 과정은, 결국 실제 가명처리에 관여하는 자 특히 '추가 정보'를 생성하거나 관리하는 자에 대한 통제를 어떻게 할 것인지가 핵심일 수 있다. 가명처리 과정의 신뢰성 확보를 위해 '신뢰받는 제3자(Trusted Third

65 회계감사기준 감사기준서 315 문단 A53-54; ISA 315 문단 부록 3 문단 22-23 참조.

66 주식회사 등의 외부감사에 관한 법률 제4조 및 동법 시행령 제5조는 재무제표를 작성하여 회사로부터 독립된 외부의 감사인에 의한 회계감사를 받아야 하는 회사의 범위를 한정하고 있고, 동법 제9조 내지 제12조 및 동법 시행령 제10조 내지 제18조는 감사인의 자격, 선임 및 보고절차, 계속감사가 가능한 범위 등을 회사의 규모 및 상장여부에 따라 달리 정하고 있는바, 가명정보의 맥락에서 입법을 통해 외부감사제도를 설계함에 있어서도 유사한 접근방식을 고려해볼 수 있다.

Party, 이하 “TTP”)’ 개념의 정립이 중요할 수 있다.[67] TTP는 특히 가명정보의 사후적 관리에 대한 신뢰성을 확보하기 위한 방안으로 유용할 수 있다.[68]

나) 일반적인 적용 가능성

일반적으로 TTP는 물리적 공간 및 조직적 구성 측면에서 데이터처리를 하는 주체와는 독립된 존재이다.[69] 이는 개인일 수도 있고 조직내의 한 부서일 수도 있다. 또는 조직 바깥의 제3의 개인이거나 조직일 수도 있다. 어떤 경우이건 TTP는 데이터처리를 하는 주체와는 조직체계나 물리적인 측면에서 독립성이 확보되어야 하는 것이 핵심이다. 어느 정도의 수준이 되어야 물리적 및 조직적 독립성이 확보되었는지를 일률적으로 판단하기는 어렵다. 예를 들어, 데이터처리를 하는 기관 내부의 한 부서라고 하더라도 실제 조직 운영에 있어서 기능상의 독립성이 충분히 보장된다면 조직적 독립성이 결여되어 있다고 말하기 어렵다.[70]

TTP의 역할의 핵심은 식별을 가능하게 하는 가명처리 비밀 등 추가 정보에 대한 접근권한과 가명처리된 데이터에 대한 접근권한 사이의 분리이다.[71] TTP의 독립성이 확보되기 위해서는 이 두 가지 역할이 직접, 간접으로 단일한 주체에게 귀속되지 않도록 설계를 해야 한다. 가명처리 비밀에 대한 접근 권한의 부여 또한 절대적인 최소한(absolute minimum)의 인원들로 제한되어야 한다.[72] 이를 통해, 가명정보를 실제로 보유하고 처리하는 부서나 조직에서는 가명처리 비밀에 대한 접근 권한이 전혀 부여되지 않고 재식별의 가능성도 존재하지 않도록 통제되어야 한다.

67 개념상 개정 개인정보 보호법 상에 규정된 ‘전문기관’이 TTP의 역할을 수행하는 것으로 해석될 수 있다(개정 개인정보보호법 제28조의3). 또한 시행령 개정령안에 규정된 한국인터넷진흥원이나 ‘전문기관’의 역할도 TTP로서의 역할이라고 해석될 수 있다. TTP의 구체적인 역할 자체도 다양하고 유연하게 설계될 수 있다.

68 International Standard Organization, “Information technology — Security techniques — Guidelines for the use and management of Trusted Third Party services”(ISO/IEC TR 14516), (2002), p.1.

69 Rolf Schwartmann & Steffen Weiß (Ed.), “Requirements for the use of pseudonymization solutions in compliance with data protection regulations”(A working paper of the Data Protection Focus Group of the Platform Security, Protection and Trust for Society and Business at the Digital Summit), (2018). 데이터 처리를 하는 주체는 GDPR에서는 주로 data controller를 가리킨다.

70 TTP의 개념 요소로 외형상의 특성을 강조해서, trusted third party와 trusted second party로 세분화하는 설명도 있다(2020. 4. 29. 방문) https://www.health-ri.nl/anonymisation-and-pseudonymisation 참조.

71 Rolf Schwartmann & Steffen Weiß (Ed.), id.

72 Rolf Schwartmann & Steffen Weiß (Ed.), id.

TTP 기능의 실제 운용에 있어서는 다양한 방식이 고려될 수 있다. 예를 들어, 식별자에 대해서는 PPT에게 접근권한이 주어지지만, 속성정보(attributes)에 대해서는 접근권한이 부여하지 않는 방식으로 운영될 수도 있다.[73] 이와는 달리, TTP를 통해 속성정보가 이동할 수 있도록 하면서 데이터베이스에 부분적으로 또는 전체적으로 암호를 적용하는 방식도 있을 수 있다.

데이터 관리의 독립성 확립에 대한 원칙으로 2인 원칙(four eyes principle, two-man rule)이 있다. 이 원칙은 어떤 조직 안에서의 특정인의 행동은 독립적인 지위에 있는 다른 사람에 의해 통제되고 확인되어야 한다는 것을 의미한다.[74] 이 원칙을 TTP에 적용하면 암호키, 가명정보 및 해당 원본 데이터에 대한 직접 또는 간접적인 접근권을 동일한 사람이 가질 수 없도록 가명처리 시스템이 설계되고 운영되어야 하는 것이 가장 중요한 기본 원칙이 된다. 만약 이 모든 권한을 궁극적으로 단일한 특정인이 독점하는 것이 가능해 진다면 개인정보가 가명처리 되기 이전의 상황으로 복원되는 것이 아렵지 않게 가능해 지고 재식별이 손쉽게 이루어질 수 있을 것이기 때문이다.

다) TTP의 적용 사례 : 어니스트 브로커(honest broker)

의료 영역에서의 데이터 통합관리의 맥락에서 도입된 어니스트 브로커(honest broker)는 TTP의 일종으로 볼 수 있다. 일반적으로 어니스트 브로커는 병원 등 조직 내에서 서로 다른 데이터베이스 시스템 사이의 중개 역할 그리고 개인정보의 이동과 정보를 관리하는 역할을 하는 개인이나 하부조직을 가리킨다.[75] 해외에서 어니스트 브로커 개념이 의료 맥락에서 본격적으로 논의된 것은 미국이 보건의료정보를 규율하는 연방법인 HIPAA(Health and Insurance Portability and Accountability Act)를 도입하면서부터이다.[76] HIPAA의 적용 대상이 되는 정보를 활용하려는 미국의 의료기관들은 환자 개인의 사전적인 동의가 없어도 활용할 수 있도록 하는 HIPAA의 비식별처리 방법을 적용하고자 했다. 이에 미국의 의료기관은 데이터를 전문적으로 수집하고 처리해서 관리하고 제공하는 조직을 도입하고 비식별처리를 하여 데이터 활용을 하고자

73 Rolf Schwartmann & Steffen Weiß (Ed.), id.

74 (2020. 4. 29. 방문) https://www.openriskmanual.org/wiki/Four_Eyes_Principle#cite_note-1 참조.

75 Boyd AD, Hosner C, Hunscher DA, Athey BD, Clauw DJ, Green LA, “An 'honest broker' mechanism to maintain privacy for patient care and academic medical research,” Int J Med Inform. pp.408 (2007).

76 Dhir, R., Patel, A. A., Winters, S., Bisceglia, M., Swanson, D., Aamodt, R., & Becich, M. J. “A multidisciplinary approach to honest broker services for tissue banks and clinical data: a pragmatic and practical model. Cancer,” 113(7), p.1710, (2008).

시도하였다.[77] 이 과정에서 어니스트 브로커의 개념이 개발되었다. HIPAA에는 비식별처리된 데이터의 경우에 일정한 조건을 충족하면 연구목적 등을 위해 활용할 수 있는 방법이 제시되어 있다. 이 때 프라이버시 침해의 위험성을 최소화하는 동시에 데이터에 대한 활용가능성을 확보하기 위한 방법으로, 의료기관 등 환자의 정보를 보유하고 있는 기관들이 어니스트 브로커라는 개념을 활용하기 시작하였다.

실무적으로 어니스트 브로커는 흔히 비식별 또는 가명처리를 수행하거나 가명처리 비밀을 보관하는 업무를 담당하고, 비식별 또는 가명 처리된 데이터를 해당 데이터를 활용하려는 연구팀에 제공하는 역할을 한다.[78] 그런 점에서, 어니스트 브로커는 정보를 제공하는 자와 정보를 이용하는 자 사이를 이어주는 중요한 역할을 담당한다. 예를 들어 어떤 연구자가 특정 유형의 데이터가 필요하다고 요청하면, 어니스트 브로커는 정보제공자도 이용자도 아닌 제3자의 관점에서, 가명처리와 정보 보안을 위한 조치를 수행하고 해당 데이터를 제공하는 역할을 한다. 어니스트 브로커의 독립성이 중요하기 때문에 어니스트 브로커는 데이터를 활용하려는 연구팀에 소속되어서는 안되고 해당 연구팀의 지휘나 명령을 받는 위치에 있어서도 안된다.[79] 그러므로 어니스트 브로커는 데이터 수탁자(trustee)의 지위에 있으면서, 환자의 식별 정보가 수집되고 저장되는 치료 영역과 비식별 또는 가명처리된 데이터를 활용하려는 연구 영역 사이에서 '장막'으로서의 역할을 해야 한다.[80]

라) TTP의 적용의 과제

어니스트 브로커와 같은 TTP가 데이터의 가명처리, 제공, 관리의 단계를 관리하는 방식이 유용하게 작동하기 위해서는, TTP 자체에 대한 신뢰성 확보가 핵심적인 관건이 된다. 특히 구체적인 방식에 따라서는 추가 정보와 가명정보 모두에 대해 TTP가 보관하거나 접근할 수 있는 상황이 일시적으로라도 발생할 수 있기 때문에, TTP를 매개로 한 내부적인 프라이버시 침해의 위험성이 존재하게 된다. 따라서 TTP의 독립성을 확보하고 이를 통해 신뢰를 구축하는 것이 중요하다.

이런 측면에서 보안의 영역에서 일반적으로 TTP에게 요구되는 다음과 같은

77 UPMC(US), "Honest Broker Certification Process Related to the De-identification of Health Information for Research and Other Duties/Requirements of an Honest Broker", Policy and Procedure Manual, (2007).

78 US UPMC Policy and Procedure Manual, Policy: HS-EC1807, (2007).

79 (2020. 4. 29. 방문) https://www.irb.pitt.edu/honest-broker-guidance 참조.

80 Dhir, R., Patel, A. A., Winters, S., Bisceglia, M., Swanson, D., Aamodt, R., & Becich, M. J., "A multidisciplinary approach to honest broker services for tissue banks and clinical data: a pragmatic and practical model." Cancer, 113(7), p.1708, (2008).

사항들이 가명처리의 맥락에서도 대체로 유사하게 적용된다고 볼 수 있다.[81] ① 적절한 내부규정을 마련한다. ② 침해 발생 시 정확한 수칙과 절차를 적용하여 대응한다. ③ 데이터 처리는 정확하게 이행하고, 명료하게 제시된 역할과 책임에 대한 규정을 준수한다. ④ 데이터와 관련된 당사자들 사이의 의사소통을 위한 절차와 인터페이스를 적절하게 마련하고 적용한다. ⑤ 관리부서 및 직원들이 규정을 준수하고 공시된 수준의 신뢰성을 달성하도록 한다. ⑥ 데이터와 관련된 절차, 운영 및 가이드라인 등 내부 수칙들에 대해 적절한 인증을 받는다. ⑦ 사용자와의 계약에 따른 의무를 준수한다. ⑧ 보안 등 문제 발생 시의 책임(liability) 가능성에 대해 명료하게 이해하고 인정한다. ⑨ 관련 법령을 준수하고, 이를 위한 관리절차를 이행한다. ⑩ 알려진 보안 위협과 그 위협에 대한 안전장치들을 명확하게 파악한다. ⑪ 위협 및 위험성에 대한 평가를 적용하고 정기적으로 갱신한다. ⑫ 적절한 조직 및 개인 차원의 방책들을 준수한다. ⑬ TTP에 대한 신뢰성을 지속적으로 확인한다. ⑭ TTP는 정부 기관에 의해 감독을 받는다.

4. 가명정보의 결합

가. 총설

개정 개인정보 보호법은 정보주체의 동의 없이도 통계작성, 과학적 연구, 공익적 기록보존 등을 위해서 서로 다른 개인정보처리자 사이에 가명정보를 결합할 수 있도록 명시적으로 규정하였다. 다만, 가명정보의 결합은 개인정보보호위원회 또는 관계 중앙행정기관의 장이 지정하는 전문기관을 통해서만 가능하도록 정하였다(개정 개인정보 보호법 제28조의3 제1항, 제28조의2 제1항). 그리고 결합된 정보를 전문기관 외부로 반출하기 위해서는 전문기관의 장의 승인을 받도록 하였다(동법 제28조의3 제2항). 개인정보 보호법 개정 전에도 2016년 개인정보 비식별 조치 가이드라인[82]은 전문기관을 통하여 비식별화 알고리즘을 활용한 정보집합물(개인정보 데이터베이스) 결합에 관한 내용을 포함하고 있었으나, 그 적법성과 실효성에 관하여 논란이 있었다.

이에 개정 개인정보 보호법은 법률 차원에서 정보의 결합에 관한 내용을 규정하였다. 그 핵심적인 내용은, ① 가명정보에 한하여 ② 전문기관을 통해서만 결합이 가능하도록 하고 ③ 결합된 정보의 반출에도 일정한 제약을 가함으로써 통제장치를 둔 것으로 요약될 수 있다. 전문기관의 개념이나 이를 통한 정보 결합의 개념은 GDPR에는 없는 것이고, 유럽 이외의 다른 나라에서도 찾아보기 어려운 것이다. 개정 개인정보 보호법은 결합 절차와 방법, 전문기관의 지정과

81 International Standard Organization, "Information technology — Security techniques — Guidelines for the use and management of Trusted Third Party services"(ISO/IEC TR 14516), (2002), p.3.

82 국무조정실 외, "개인정보 비식별 조치 가이드라인" (2016).

지정 취소 기준·절차, 관리·감독, 결합된 정보의 반출 및 승인 기준·절차 등의 구체적 내용을 대통령령에 위임하고 있고(동법 제28조의3 제3항), 동법 시행령 일부개정령안은 그 내용의 일부를 정한 뒤 나머지는 개인정보 보호위원회가 정하여 고시하도록 재차 위임하고 있다(동안 제29조의2 제6항, 제29조의3 제6항).

개정 개인정보 보호법에 의하여 도입된 가명정보 결합 제도에 관하여는 향후 실무에 있어 복잡다단하면서도 중요한 쟁점들이 다수 등장할 것으로 예상된다. 그 논의의 구체적인 방향은 후술하는 바와 같이[83] 향후 고시 등 하위법령 및 가이드라인에 구체적으로 어떠한 내용이 담기느냐에 따라 크게 달라질 것이어서, 이 글에서는 향후 논의를 위한 전제로서 ① '서로 다른' 개인정보처리자 간 가명정보 '결합'의 의미, ② 가명정보 결합의 문제상황, ③ 가명정보 결합의 실익, ④ 가명정보 결합에 있어 전문기관의 지위와 역할의 네 가지 핵심 쟁점들을 제시하고, 이에 대하여 간략히 살펴보기로 한다.

나. 가명정보 결합의 주요 쟁점들

1) '서로 다른' 개인정보처리자 간 가명정보 '결합'의 의미

개정 개인정보 보호법은 '결합'의 개념을 별도로 규정하고 있지 않다. 따라서 가명정보의 결합이 무엇을 의미하는지 실무적, 이론적으로 문제될 수 있다. 대부분의 개인정보가 구조화된 데이터(structured data)로서 통상 행과 열로 구성된 하나의 표 형태로 표현가능하다고 가정할 경우에, 제한 없이 이를 '합쳐(combine, consolidate)' 더 큰 형태의 표를 만드는 것만이 결합에 해당하는 것인지, 아니면 서로 다른 개인정보처리자 간 가명정보를 상호 비교하는 등 일종의 '연결(link)' 행위까지 결합에 해당하여, 이런 유형의 작업을 모두 전문기관을 통해서만 진행해야 하는 것인지 명확하지 않다. 구조화되지 않은 데이터의 경우에도 그 결합의 개념은 명확하지 않다. 또한 서로 다른 유형의 데이터 사이의 결합이 무엇을 의미하는지, 가령 영상의 형태로 된 가명정보와 문자열의 형태로 된(텍스트) 가명정보를 결합한다는 것은 무엇을 의미하는지도 명확하지 않다.[84]

한편 개정 개인정보 보호법은 '서로 다른 개인정보처리자 간' 가명정보의 결합을 규정하고 있는데, 그렇다면 예를 들어 동일한 기업집단에 속하는 두 계열회사가 각기 보유하는 가명정보를 결합하는 경우에도 전문기관을 통해야만 하는지 등 '서로 다른 개인정보처리자'의 기준을 어떻게 정할 것인지도 문제될 수 있다. 어렵지 않게 생각할 수 있는 하나의 기준으로 '법인'을 생각할 수 있는데, 왜

83 특히 가명정보 결합에 있어 전문기관의 지위와 역할에 관한 Ⅳ. 2. 라.의 설명 참조.

84 가령 특정 지역 내 코로나19 확진자의 동선이 일부 담긴 CCTV 영상을 모자이크 등을 활용하여 가명처리하고 해당 지역 내 영업장에서의 거래내역 표를 가명처리하여 결합하는 경우를 생각해 볼 수 있다. 실제로 의료영역에서는 다양한 유형의 영상정보가 생성된다.

법인이 기준이 되어야 하는지는 명확하지 않다.

2) 가명정보 결합과 관련된 리스크

어떠한 방식으로든 서로 다른 두 개인정보 데이터베이스를 '결합'하면 한편으로는 통계적 분석 등에 있어서의 활용도가 증대되지만, 다른 한편으로는 정보량이 늘어남으로 인해 정보주체가 재식별되거나 프라이버시가 침해될 위험이 증가하게 되는 것이 일반적이다.[85] 유용성 제고와 개인정보 보호가 상충하는 문제가 개인정보의 결합에 있어서도 마찬가지로 발생하게 되는 것이다. 따라서 개정 개인정보 보호법에 따른 가명정보의 결합과 관련하여 제기될 수 있는 다양한 세부 쟁점들을 논함에 있어서는 이러한 리스크 증대 등을 면밀히 고려해야 한다.

그런데 정보주체가 재식별되거나 프라이버시가 침해될 위험에 관해서는, 이를 야기하는 주체에 따라 ⓐ 해당 가명정보를 보유하는 개인정보처리자에 의한 위험, ⓑ 결합전문기관에 의한 위험, ⓒ 해커 등 잠재적인 제3의 공격자(motivated intruder)[86]에 의한 위험으로 나누어서 생각할 수 있다. 이 경우 통상 논의의 대상이 되는 것은 ⓐ 또는 ⓒ이나, 가명정보 결합의 맥락에서는 ⓑ를 함께 고려해야 한다. 가명정보의 결합을 수행하는 전문기관은 앞서 살펴본 것과 같은 TTP로서 도입된 것이기는 하지만, 실제로 결합을 수행하는 부서와 결합에 필요한 결합키 등 각종 비밀을 기록·보관하는 부서를 분리하여 운영하는 등으로 결합의 전 과정에 거쳐 적절한 기술적·관리적 보호조치가 이루어지지 않는다면 전문기관 내부에서 정보주체가 재식별되거나 그 프라이버시가 침해될 가능성이 있기 때문이다.

3) 가명정보 결합의 실익

가명정보의 결합은 무시할 수 없는 수준의 비용을 수반하게 된다. 결합과 관련된 다양한 환경의 구축은 물론 필요한 통제장치들을 마련하고 이행하는 데에 상당한 인적, 물적 자원이 소요될 수밖에 없기 때문이다. 그렇다면 개인정보처리자가 가명정보를 결합함으로써 얻을 수 있는 편익은 무엇인가?

가장 큰 잠재적 편익으로 생각할 수 있는 것은, 고객의 동의 없이도 가명정보를 결합함으로써 상세한 분석을 하여(analytics 또는 profiling) 이를 토대로 맞춤형 서비스를 제공(targeted service)하고 맞춤형 마케팅(targeted marketing)을 실시하는 것이다. 하지만 개정 개인정보 보호법상 이러한

85 이동진, "데이터거래의 법적 쟁점 및 데이터거래 가이드라인", 서울대학교 인공지능정책 이니셔티브 이슈페이퍼 (2019. 11.), 30면.

86 이에 관한 보다 상세한 설명으로는 이동진, "개인정보 보호법 제18조 제2항 제4호, 비식별화, 비재산적 손해 - 이른바 약학정보원 사건을 계기로 -", 정보법학 제21권 제3호 (2017. 12.), 253-284면 중 266-268면 참조.

행위가 허용되는 것인지, 허용된다면 어떤 수준으로 허용되는지는 불분명하다. 우선 개인 맞춤형 서비스의 제공이나 마케팅 활동을 위한 가명정보의 결합이 '통계작성, 과학적 연구, 공익적 기록보존' 등을 위한 것에 해당하는지 여부가 문제될 수 있다(개정 개인정보 보호법 제28조의3 제1항). 또한 가명정보의 결합은 그 자체로 개정 개인정보 보호법상 '처리'에 해당하는데, "누구든지 특정 개인을 알아보기 위한 목적으로 가명정보를 처리"하는 것은 금지될 뿐더러 "개인정보처리자는 가명정보를 처리하는 과정에서 특정 개인을 알아볼 수 있는 정보가 생성된 경우에는 즉시 해당 정보의 처리를 중지하고, 지체 없이 회수·파기" 해야 하는 의무가 있기도 하다(개정 개인정보 보호법 제28조의5 제1항, 제2항). 결국 개인 맞춤형 서비스나 마케팅을 제공하기 위한 목적으로 가명정보를 결합하려 한다면, '특정 개인을 알아본다'는 것의 정확한 의미가 무엇인지에 관한 논의가 선행되어야 한다.

다만 개인을 특정하거나 재식별하는 대신, 가령 결합된 가명정보 데이터베이스 '전체'를 통계적으로 분석하는 것은 어렵지 않게 생각할 수 있다(개정 개인정보 보호법 제28조의3 제1항, 제28조의2 제1항). 일례로, 구두를 판매하는 A회사와 가방을 판매하는 B회사가 각기 보유한 고객별 구매내역 데이터베이스를 서로 결합하는 경우를 생각해보도록 하자. 앞서 본 바와 같이 A회사와 B회사가 결합 데이터베이스로부터 특정 고객을 재식별 해내어 그가 관심을 가질 만한 상품의 할인 쿠폰을 발송하는 등의 개인 맞춤형 마케팅을 실시하는 것이 허용되는지는 불명확하다. 반면, 결합 데이터베이스 전체를 통계적으로 분석하여 "X구두를 사는 고객층과 Y가방을 사는 고객층은 상당 부분 중복된다"와 같은 특징을 발견해내는 것은 고객들의 동의 없이도 가능할 것으로 보인다. A회사와 B회사는 이러한 발견을 토대로 X구두와 Y가방 중 하나를 구매하는 모든 고객들에게 - 개별 고객을 특정할 필요 없이 - 다른 상품의 가격을 할인하여 주는 판촉활동을 진행할 수도 있을 것이다.

4) 가명정보 결합에 있어 전문기관의 지위와 역할

가명정보의 결합과 관련하여 향후 실무적으로는 전문기관의 지위와 역할을 둘러싼 논란이 발생할 가능성이 있다. 가령 개인정보처리자들이 가명정보의 결합을 신청한 경우 전문기관이 이를 스스로의 판단 하에 거절하거나 반려할 수 있는지 불분명하다. 만약 개인정보처리자들이 그 결합의 절차나 방법 등을 구체적으로 정하여 신청하였다면 전문기관은 특별한 사정이 없는 한 이에 따라야 하는 것인지도 문제될 수 있다.

전문기관의 지위와 역할에 관하여는 크게 두 가지 모델을 상정해볼 수 있다. 우선 하나의 모델(이하 '소극적 전문기관 모델')은 전문기관의 역할은 각 개인정보처리자로부터 제공받은 가명정보를 단순히 결합하는 데 국한되며 결합된 가명정보에 따른 각종 위험은 기본적으로 결합을 신청한 개인정보처리자가 부담하여야 한다는 입장에 기초한 모델이다. 반면 다른 하나의 모델(이하 '적극적 전문기관 모델')에 따르면, 전문기관은 가명정보의 결합에 따라 발생할 수 있는 각종 위험에 대응하기 위한 각종 보호조치를 결합의 전 과정에 걸쳐 스스로의 판단 하에 합리적인 범위 내에서 취할 것이 요구된다. 이는 전문기관의 통제장치로서의 기능을 강조하는 입장이다. 두 모델 사이에서 어떤 선택을 하느냐에 따라 개정 개인정보 보호법이 하위 법령에 위임한 ① 가명정보 결합의 절차와 방법(특히 결합된 정보의 반출 방법), ② 전문기관의 지정 및 지정취소 기준(특히 순수 민간기관을 전문기관으로 지정할 것인지 여부), ③ 전문기관에 대한 관리·감독 등에 관한 구체적 내용이 달라질 것이다.

개정 개인정보 보호법의 문언만으로는 개정법이 위 두 모델 사이에서 어떠한 입장을 취하고 있는지 분명하지 아니하나, 현재의 시행령 일부개정령안은 적극적 전문기관 모델에 가까운 입장을 취하고 있는 것으로 보인다. 예를 들어 시행령 일부개정령안 제29조의2 제2항은 전문기관으로 하여금 결합신청자나 전문기관 스스로의 독자적 판단이 아닌, "특정 개인을 알아볼 수 없도록 [개인정보] 보호위원회가 정하여 고시하는 절차와 방법에 따라 가명정보를 결합"하도록 규정하고 있기 때문이다. 또한 시행령 일부개정령안은 결합신청자가 "결합전문기관에 설치된 안전성 확보에 필요한 기술적·관리적·물리적 조치가 된 공간", 즉 전문기관에 의해 정해진 "분석공간"에서 결합된 정보를 분석하는 것을 원칙으로 정하고 있기도 하다(시행령 일부개정령안 제29조의2 제3항 및 제4항).[87]

5. 결론

개인정보를 안전하게 보호하면서 합리적인 범위 내에서 데이터 활용의 기반을 조성하기 위하여 데이터 3법이 개정되었으며, 개정 내용의 핵심 중 하나는 가명정보와 관련된 것이다. 가명정보 특례를 둘러싸고 매우 다양한 쟁점이 남아있다. 본 연구는 가명정보 및 가명처리의 개념에 대해 살펴보고 관련된 쟁점들을 분석하고 정리하였다.

가명정보는 원래의 상태로 복원하기 위한 추가 정보의 사용·결합 없이는 특정 개인을 알아볼 수 없는 정보인 것으로 규정된다. 여전히 개인정보로 분류되지만, 별도 유형의 특수한 개인정보로 규정되어 정보주체의 동의 없이 통계 목적, 과학적 연구 목적, 공익적 기록보존 목적 등을 달성하기 위하여 이용하거나 제공할 수 있다. 또한 개정법은 가명정보의 결합에 대한 규정을 도입하고 있는데, 결합은 전문기관을 통해 수행해야 하는 것으로 규정되었다.

데이터 3법의 개정을 통해 도입된 변화는 이론적으로나 실무적으로나 적지 않은 새로운 과제를 제시한다. 여러 관련 쟁점들에 대한 상세하고 진지한 논의를 통해 데이터 시대로의 변화에 적극적으로 대응할 필요가 있다.

87 이와 달리 2020. 3. 31. 입법예고된 신용정보법 시행령 일부개정령안 제14조의2 제3항 제3호 및 제5호는 결합을 수행한 전문기관으로 하여금 원칙적으로 그 결과물을 결합의뢰기관에게 '전달'하도록 규정하고 있다.

참고문헌

- 국무조정실 외, “개인정보 비식별 조치 가이드라인” (2016).
- 관계부처 합동 브리핑, 개인정보 보호법 개정 후속조치 계획, (2020. 1. 21.).
- 김한겸, 유전자은행 업무지침 개발연구, 보건복지부 (2005).
- 이동진, “개인정보 보호법 제18조 제2항 제4호, 비식별화, 비재산적 손해 - 이른바 약학정보원 사건을 계기로 -”, 정보법학 제21권 제3호 (2017. 12.).
- 이동진, “데이터거래의 법적 쟁점 및 데이터거래 가이드라인”, 서울대학교 인공지능정책 이니셔티브 이슈페이퍼 (2019. 11.).
- 이동진, “목적합치의 원칙과 가명정보의 특례”, 법률신문 연구논단, (2020. 3. 23.).
- 이창우·송혁준·전규안·권오상, 회계감사 Study Guide (제6판), 경문사 (2019).
- 참여연대, “[이슈리포트] 그 많은 내 개인정보는 누가 다 가져갔을까 - 2007-2017 개인정보수난사 Worst 44”, (2018. 10. 1.).
- 행정안전부, 데이터 규제 혁신, 청사진이 나왔다, 보도자료, (2018. 11. 22.).

- 헌법재판소 2005. 5. 26. 선고 99헌마513 결정.
- 헌법재판소 2007. 5. 31. 선고 2006헌가10 전원재판부 결정.
- 대전지방법원 논산지원 2013. 8. 9. 선고 2013고단17 판결.
- 서울중앙지방법원 2011. 2. 23. 선고 2010고단 5343 판결.
- 서울중앙지방법원 2017. 9. 11. 선고 2014가합508066 판결.
- 서울중앙지방법원 2019. 5. 3. 선고 2017나2074963 판결.
- 서울중앙지방법원 2020. 2. 14. 선고 2015고합665 판결.

- Article 29 Data Protection Working Party, "Opinion 4/2007 on the concept of personal data" (2007).
- Article 29 Data Protection Working Party, "Opinion 01/2012 on the data protection reform proposals" (2012).
- Article 29 Data Protection Working Party, "Opinion 05/2014 on Anonymisation Techniques" (2015).
- Arvind Narayanan and Vitaly Shmatikov, Robust De-anonymization of Large Sparse Datasets, IEEE Symposium on Security and Privacy (2008).
- Bourka, Athena, Prokopios Drogkaris, European Union, and Agency for Network and Information Security. “Recommendations on Shaping Technology According to GDPR Provisions: An Overview on Data Pseudonymisation” (2018). http://dx.publications.europa.eu/10.2824/74954.
- Bourka, Athena, Prokopios Drogkaris, Ioannis Agrafiotis, and European Network and Information Security Agency. “Pseudonymisation Techniques and Best Practices: Recommendations on Shaping Technology According to Data Protection and Privacy Provisions” (2019). https://op.europa.eu/publication/manifestation_identifier/PUB_TP0119810ENN.
- Boyd AD, Hosner C, Hunscher DA, Athey BD, Clauw DJ, Green LA, “An 'honest broker' mechanism to maintain privacy for patient care and academic medical research,” Int J Med Inform (2007).
- Choi, H. J., Lee, M. J., Choi, C. M., Lee, J., Shin, S. Y., Lyu, Y., Park, Y. R., & Yoo, S., Establishing the role of honest broker: bridging the gap between protecting personal health data and clinical research efficiency. PeerJ, 3 (2015).
- Committee of Sponsoring Organizations of the Treadway Commission(COSO), Internal Control - Integrated Framework, Executive Summary (2013. 3.).
- Cynthia Dwork Moni Naor Toniann Pitassi Guy Rothblum, "Differential Privacy Under Continual Observation", STOC '10: Proceedings of the 42nd ACM symposium on Theory of computing (June 2010).
- Cynthia Dwork, Adam Smith, Thomas Steinke, and Jonathan Ullman, Exposed! A Survey of Attacks on Private Data, Annual Review of Statistics and Its Application Vol. 4:61-84 (March 2017).
- Dhir, R., Patel, A. A., Winters, S., Bisceglia, M., Swanson, D., Aamodt, R., & Becich, M. J. “A multidisciplinary approach to honest broker services for tissue banks and clinical data: a pragmatic and practical model. Cancer,” 113(7) (2008).
- European Commission, "Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)", (2012).
- El Emam K, Rodgers S, Malin B. "Anonymising and sharing individual patient data.", PMC, (2015).
- ENISA, “Pseudonymisation Techniques and Best Practices”, (2019).
- ENISA, “Recommendations on Shaping Technology according to GDPR Provisions” (2018. 11.).
- International Standard Organization, “Information technology — Security techniques — Guidelines for the use and management of Trusted Third Party services”(ISO/IEC TR 14516:) (2002).
- International Standard Organization, “Information technology — Security techniques — Guidelines for the use and management of Trusted Third Party services”(ISO/IEC TR 14516:) (2002).
- L. Sweeney. k-anonymity: a model for protecting privacy. International Journal on Uncertainty, Fuzziness and Knowledge-based Systems, 10 (5), (2002).
- Ohm, Paul, Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization, UCLA Law Review, Vol. 57, (August 13, 2009).
- Protection Focus Group of the Platform Security, Protection and Trust for Society and Business at the Digital Summit (2018).
- Rolf Schwartmann & Steffen Weiß (Ed.), “Requirements for the use of pseudonymisation solutions in compliance with data protection regulations”(A working paper of the Data Protection Focus Group of the Platform Security, Protection and Trust for Society and Business at the Digital Summit), (2018).
- UPMC(US), “Honest Broker Certification Process Related to the De-identification of Health Information for Research and Other Duties/Requirements of an Honest Broker”, Policy and Procedure Manual, (2007).
- UPMC(US), Policy and Procedure Manual, Policy: HS-EC1807, (2007).

서울대학교
인공지능정책
이니셔티브 안내

서울대학교 인공지능정책 이니셔티브는 인공지능과 관련된 다양한 사회경제적, 법적, 정책적 이슈들을 연구하고 논의하기 위해 시작된 서울대학교 법과경제연구센터의 프로그램입니다. '소셜랩(Social Lab)' 개념을 지향하여, 여러 배경과 관심을 가진 분들 사이의 협업과 지속적인 대화를 추구합니다. 서울대학교 법학전문대학원의 고학수 교수와 임용 교수가 함께 이끌고 있습니다.

1. 발간문 안내

서울대학교 인공지능정책 이니셔티브의 주요 발간물은 이슈페이퍼와 워킹페이퍼가 있고, 비정기적으로 발간되는 단행본 및 학술행사 자료집 등이 있습니다. 이슈페이퍼와 워킹페이퍼 등의 자료들은 홈페이지를 통해 다운로드 받으실 수 있습니다.

2. 행사 안내

서울대학교 인공지능정책 이니셔티브의 주요 행사는 이슈페이퍼를 발표하고 논의하는 행사(상반기 및 하반기 각 1회) 그리고 국내외 연구자들을 초빙하여 진행하는 대규모 국제학술대회(연 1회) 등이 있습니다. 그 이외에 비정기적으로 진행하는 행사들도 있습니다.

3. 이슈페이퍼 2020-1

이번 이슈페이퍼는 서울대학교 인공지능정책 이니셔티브의 세 번째 이슈페이퍼로, 2020. 05. 07.에 열린 웨비나 행사를 위해 준비되었습니다.

http://sapi.co.kr/