

개정된 데이터 3법에 따른 가명정보 활용 방법 및 처리 시 보안 주의사항

■ 가명정보의 배경 및 취지

가명정보는 개인정보인가?



지난 20년 8월, 홍수처럼 밀려드는 데이터에 대한 활용 및 규제 방안 등을 주제로 논의가 이뤄졌다. 그 결과 “개인정보보호법”, “신용정보의 이용 및 보호에 관한 법률(이하, 신용정보법)”, “정보통신망 이용촉진 및 정보보호 등에 관한 법률”이 데이터 3법이라는 주제로 개정되었다.

당시 개인정보보호법 및 신용정보법의 입법 취지에도 명시되어 있듯이 가명정보¹는 마이데이터(본인신용정보관리업)와 더불어 데이터 3법의 핵심내용이다.

¹ 가명정보: 이름, 주민등록번호와 같은 개인정보를 가명 처리함으로써 원래의 상태로 복원하기 위한 추가 정보의 사용·결합 없이 특정 개인을 알아볼 수 없는 정보

이는 법으로 데이터 활용과 개인정보보호라는 상호 반비례하는 사안에 대한 적절한 중간지점을 제시함으로써 “개인정보를 가명정보로 활용하려면 법적 요건들을 준수해라”라는 의미로 해석할 수 있다.

법령에서 말하고 있는 가명정보에 대한 주요 요건은 아래와 같이 크게 세 가지로 볼 수 있다.

첫 번째는 가명정보를 처리하고자 할 때는 그 목적이 과학적 연구, 통계 작성, 공익적 기록 보존 등으로 활용 때에만 정보주체의 동의 없이 처리할 수 있다.

두 번째는 가명정보는 특정 개인으로 식별되면 안 된다. 가명정보 자체는 물론 다른 정보와 결합의 가능성도 고려해서 개인 식별 여부를 검토해야 한다.

세 번째는 가명정보도 개인정보이다. 가명처리² 과정 혹은 가명정보 저장 등에 대해서도 마땅히 개인정보로 취급하여 처리해야 한다. 가명정보를 처리하는 시스템은 개인정보 처리시스템으로 분류하여 관계 법령 및 고시에서 요구하는 규정을 준수해야 한다. 해당 정보가 분실·도난·유출·위주·변조 또는 훼손되지 않도록 안전성 확보에 필요한 기술적·관리적 및 물리적 보호 조치를 수행해야 한다.

² 가명처리: 개인정보 중 전체 또는 일부를 대체하는 등과 같은 방법으로 특정 개인을 알아볼 수 없도록 처리하는 과정

■ 가명정보의 활용 사례

가명정보는 어떻게 사용되나?

데이터 3법이 개정되기 전, 기관/기업들은 개인정보 수집 시에 고지한 목적 내에서만 개인정보를 이용할 수 있고, 수집 시 고지한 목적 외로 이용/제공하려면 목적 외 이용/제공에 대해 정보주체로부터 별도 동의를 받아야 했다. 하지만 이제는 법의 준수 사항을 만족하면 가명정보를 통해 보다 넓은 데이터 활용을 할 수 있게 되었다.

기관/기업에서 가명정보를 활용하는 방법도 두 가지로 나뉘는데 내부 활용과 외부 결합이다.

내부 활용은 기관/기업 자체로 보유하고 있는 개인정보를 활용하는 것으로 범위/재식별성을 고려하여 처리할 수 있다. 데이터 3법이 개정되기 전에 기관/기업들은 내부에 보유하고 있던 개인정보를 통해 통계 및 연구 업무를 수행하고 있었다.

외부 결합은 각 기관/기업들이 소유한 개인정보가 가명 처리되어 상호 결합되는 것을 일컫는다. 이를 통해 생성된 결합 정보는 다양한 Insight를 얻을 수 있는 자료가 되고 그에 따라서 새로운 비즈니스가 도출될 수 있다.

이런 상황을 반영하듯 최근 주요 가명 처리에 대한 사례를 보면, 각 기관/기업들이 그동안 다루기 어려웠던 분야 또는 보유하지 못한 정보를 결합하여 정부기관의 정책 추진뿐만 아니라 기업 간 이윤창출을 위한 가명정보 결합이 이루어지고 있고, 그 수요가 점차 많아지고 있다.

infosec

사례	정보보유 기관	결합정보	이용기관	결합전문기관
1인 가구 복지정책 개발	A지자체	가구정보, 소득정보, 직업 및 주택정보	A지자체	ㄱ결합기관
	B기업	위치정보, 통신 이용량		
친환경차 충전 인프라 수요 조사 및 입지 선정	C지자체	등록 차량 정보	C지자체	ㄴ결합기관
	D기업	내비게이션 주행 정보		
지역/연령 등에 따른 구매 패턴 분석	E편의점	지역정보, F카드를 통한 거래 및 품목 정보	F카드사 E편의점	ㄷ결합기관
	F카드사	카드 소유주 정보		

* 출처: 개인정보보호위원회 2021 가명정보 활용 우수사례집

■ 가명정보의 보호

가명정보 처리 시 주의사항은?

가명정보의 활용 사례에서 보듯이 이미 우리는 알게 모르게 가명정보처리³를 통한 정책 및 서비스 등을 직/간접적으로 제공받고 있다. 이렇게 다양하게 활용되는 만큼 가명정보의 보호에도 신경 써야 한다.

가명정보보호의 특징 중 하나로 개인정보로의 재식별을 꼽을 수 있다. 앞서 이야기했듯이 가명정보는 개인정보로 식별되지 않게 가명 처리 후 활용이 가능하다. 가명정보가 개인정보로 재식별이 된다면 이 가명정보는 즉시 파기 또는 회수하도록 법에서 요구하고 있다.

이는 가명처리가 미흡하게 되어 가명정보에서 바로 개인정보가 식별될 수도 있지만, 아래처럼 적절하게 가명처리가 되었다라든가 가명정보 취급자나 가명정보처리시스템의 운영 환경에 따라서도 재식별 가능성이 존재할 수 있기 때문이다.

먼저, 가명처리가 적절하게 잘 처리되어 재식별이 어렵더라도 남들이 모르는 특정 정보를 알고 있는 가명정보 취급자가 해당 가명정보를 처리한다면 재식별의 경우가 발생할 수 있다.

예를 들어 아래와 같이 서울 소재 한 백화점의 9월 매출에 대해서 가명 처리된 통계 내역이 있다고 가정하자.

infosec

구분	이름	생년월일	A매장(의류) 지출	B매장(가전) 지출	...	Z매장(식당) 지출
처리방법	이름 삭제	연령대로 치환	만 단위 반올림	만 단위 반올림	...	만 단위 반올림
Field	홍OO	40대	9,200,000	13,570,000	...	100,000

해당 정보는 이름 없이 성과 40대라는 연령, 지출액 만으로 특정 개인을 식별하는 것이 가능해 보이지 않는다.

하지만, 해당 정보를 A매장 직원이 보면 9월에 920만 정도 결제한 40대 홍 씨는 충분히 누구인지 식별이 가능해 보인다. 또한, 만약 해당 정보주체가 백화점 멤버십을 적립했다면 결제 금액에 따른 멤버십 적립/소진 정보는 멤버십 담당자가 기존에 알고 있는 정보이고, 해당 멤버십 정보와 가명정보를 비교하면 멤버십 담당자에게는 해당 가명정보는 개인 식별이 가능한 정보가 된다.

결국 해당 가명정보 자체도 중요하지만 ‘가명정보를 누가 취급하냐’도 매우 중요하다. 이런 재식별의 가능성을 줄여보고자 가명정보 취급자와 개인정보 취급자는 직무분리/권한설정 등을 통해 접점이 없도록 해야 한다.

³ 가명정보처리 : 가명정보에 대한 이용, 제공, 결합 등의 행위

두 번째로 가명정보는 추가정보⁴ 없이는 개인을 식별할 수 없기에 법령에서는 파기에 대해 예외 조항을 두고 있다. 가명정보는 목적 달성 후 파기해야 하는 개인정보와는 다르게 파기를 하지 않아도 된다. 그렇지만, 가명정보를 활용하면서 처리한 가명정보들이 계속 누적될 수 있다. 단순히 가명정보 집합 내 건수가 누적된다면 큰 문제가 없을지도 모르지만 가명정보의 속성값이 누적된다면 재식별이 될 수도 있다.

예를 들어 21년에 A가명정보 “지역별 임신부의 연령 통계”를 처리하고 22년에 B 가명정보 “연령 대 별 임신부의 질병, 처방 약물 통계”를 처리했다고 가정하자. A와 B를 따로 보면 해당 정보로 개인을 식별하기는 어려워 보인다. 하지만 파기하지 않고 누적되어 A/B가 합쳐지고 속성값이 늘어나면 “지역/연령/임산부/질병/처방 약물”의 정보가 되어 군소도시처럼 표본이 적은 곳에서는 해당 정보로도 누구의 정보인지 재식별 가능성이 높아진다.

이렇게 가명정보의 파기가 이루어지지 않는다면 누적된 속성값으로 인해 개인을 식별할 수도 있으므로 가명정보의 목적을 달성하면 파기하는 것이 바람직하다.

마지막으로, 가명정보는 원장 개인정보와 가명처리에 사용된 추가정보와는 분리해야 한다. 추가정보는 가명처리에 사용되는 정보이므로 가명정보와 추가정보를 동시에 알게 된다면 가명정보에서 역순으로 처리해 원래 개인정보가 도출될 수 있다. 마찬가지로 가명정보와 원장 개인정보가 조합이 된다면 비슷한 정보 혹은 특이치 정보로 인해 추가정보를 추측할 수 있고 곧 나머지 정보도 개인정보로 재식별 될 수 있으므로, 세 데이터를 각각 분리보관해야 하고 접근을 통제해야 한다.

⁴ 추가정보 : 가명처리 과정에서 생성/사용된 정보 예) 알고리즘, 매칭 테이블, 암호키 등

■ 마치며

가명정보는 가명/결합정보 생성에 그 목적을 두는 것이 아니다. 앞서 몇몇 결합 사례를 본 것과 같이 결합된 가명정보로 끝나는 것이 아니라 그 결합정보를 기반으로 정책을 수립하고, 소비 패턴을 분석하여 맞춤형 서비스 제공 확대하는 등 정책 및 비즈니스 계획 수립에 중요한 자료가 될 수 있다. 기관/기업 간 사업제안이나 제휴 시에도 그 근거로써 충분히 사용될 가치가 있고, 기타 기업 간 협업 비즈니스를 시도할 때에도 사전에 해당 기업과 가명정보 결합을 통해 비즈니스 모델을 그려볼 수 있다. 이런 데이터 활용을 통해 사업의 확장뿐 아니라 비즈니스 리스크를 낮추기 위한 근거로 쓰일 수도 있다.

가명정보의 활용에서 가장 중요하게 고려해야 될 사항은 재식별이다. 가명정보만의 특이성은 이 재식별성에서 찾을 수 있고, 당연히 보안적인 시각에서도 이 재식별을 집중적으로 살펴봐야 한다.

더불어 [가명정보의 보호] 예시의 재식별 위험을 제거하지 않을 경우 개인정보법 제75조 ①항에 의거하여 과태료 처분을 받을 수도 있다.

제28조의5(가명정보 처리 시 금지의무 등) ② 개인정보처리자는 가명정보를 처리하는 과정에서 특정 개인을 알아볼 수 있는 정보가 생성된 경우에는 즉시 해당 정보의 처리를 중지하고, 지체 없이 회수·파기하여야 한다.

제75조(과태료) ① 다음 각 호의 어느 하나에 해당하는 자에게는 5천만원 이하의 과태료를 부과한다.

7의2. 제28조의5제2항을 위반하여 개인을 알아볼 수 있는 정보가 생성되었음에도 이용을 중지하지 아니하거나 이를 회수·파기하지 아니한 자

결론적으로 가명정보는 개인정보이므로 법령에서 요구하는 개인정보의 안전성 확보에 필요한 조치뿐 아니라 데이터 자체의 속성을 분석/파악하는 등 재식별을 방지하기 위한 추가적인 보호 조치가 필요하다.