금융분야 가명·익명처리 안내서

2020. 8. 6

금융위원회

금융감독원

발 간 사

2020년 2월 개인정보보호법, 신용정보법, 정보통신망법(데이터 3법) 개정안이 국회를 통과함에 따라 데이터 경제 활성화를 위한 제도적 기반이 마련되었습니다. 개정법의 시행일인 2020년 8월 5일 부터는 데이터 3법에 근거하여 금융회사, 상거래 기업 등이 가명정보, 익명정보를 안전하게 활용할 수 있게 되었습니다.

은행·카드·보험·금융투자 등 금융업권의 데이터와 통신정보·위치정보·보건의료정보 등 다른 산업분야에서 관리되고 있는 다양한 형태의 데이터를 서로 융합하여 금융분야를 비롯한 전 산업 분야의 혁신성장을 이끌 수 있게 되었습니다.

가명처리와 익명처리 및 데이터 결합 절차가 새롭게 도입됨에 따라 가명정보 및 익명정보를 실제 활용하고자 하는 금융회사와 일반 기업에서는 구체적으로 어떤 방식으로 가명·익명처리 및 데이터 결합을 수행해야 하는지에 대한 많은 질문이 있어 왔습니다.

이 안내서를 통해 안전한 가명처리 및 익명처리 방법과 데이터전문기관이 제공하는 데이터 결합 방법 등을 안내함으로써 현장의 법적 불확실성을 해소하는 한편, 가명정보 및 익명정보를 안전하게 활용하고, 결합할 수 있도록 도와드리고자 합니다.

이 안내서가 디지털 전환(디지털 트랜스포메이션)을 통한 혁신성장을 지원하고, 안전한 개인정보 활용을 통해 정보주체의 개인정보 자기결정권을 보장하는 역할을 할 수 있기를 기원합니다.

안내서 발간에 참여해주신 데이터전문기관, 금융회사, 기업, 여러 전문가 분들께 감사드립니다.

2020. 8. 6. 금융위원회·금융감독원

목 차

Ⅰ. 개요 1
1. 추진배경 및 목적
2. 용어정의
3. 개인정보, 가명정보, 익명정보
4. 금융분야 가명·익명처리 일반

Ⅱ. 가명처리 16
1. 개요
2. 가명처리 절차
3. 가명처리 방법
4. 가명처리에 관한 행위 규칙
5. 가명정보 및 추가정보에 관한 보호조치 기준

Ⅲ. 익명처리 및 적정성 평가 53
1. 개요
2. 익명처리 방법
3. 적정성 평가

Ⅳ. 정보집합물 결합 63
1. 개요
2. 정보집합물 결합 절차
3. 데이터전문기관 보유 데이터와 외부정보의 결합
4. 주기적·반복적 정보집합물 결합 및 활용

【부록 1】 가명·익명처리 기법
【부록 2】 익명처리 적정성 평가 기초자료 작성 방법(예시)
【참고문헌】

Ⅰ. 개요

1. 추진배경 및 목적

가. 추진배경

데이터 3법*이 2020년 1월 국회 본회의에서 가결됨에 따라 가명정보, 익명정보를 법에 근거하여 활용할 수 있는 길이 열렸다. 이로 인해 은행・카드・보험・금융투자 등 금융업권 별로 체계적으로 관리되는 정형데이터와 통신정보・위치정보・보건의료정보 등 다른 산업분야에서 관리되고 있는 다양한 형태의 데이터를 서로 융합하여 금융분야의 혁신 성장을 이끌 수 있게 되었다.

* 「개인정보 보호법」, 「정보통신망 이용촉진 및 정보보호 등에 관한 법률」(이하 「정보통신망법」), 「신용정보의 이용 및 보호에 관한 법률」(이하 「신용정보법」)

개정된 「신용정보법」은 금융분야에서 기존 『개인정보 비식별 조치 가이드라인』(2017.7.1.시행)의 법적 한계를 극복하고 가명・익명처리한 정보를 안전하게 활용할 수 있는 제도적 기반을 제공할 것으로 예상된다. 개정법에서는 통계작성(시장조사 등 상업적 목적의 통계작성 포함), 연구(산업적 연구 포함), 공익적 기록보존 등을 위하여 개인인 신용정보 주체의 동의 없이 개인신용정보를 가명처리 하여 사용할 수 있을(동법 제32조의2제6항 제9의2호)뿐만 아니라, 누구인지 알아 볼 수 없도록 익명처리한 경우에는 목적 제한 없이 자유로운 활용도 가능하게 되었다(동법 제40조의2제4항).

본 안내서는 개정 「신용정보법」, 동법 시행령 및 하위 규정에 따른 가명・익명처리 및 활용에 대한 하나의 예시를 제시하여, 가명・익명처리에 대한 이해도를 높이고 가명정보・익명정보의 안전한 활용을 돕기 위해 마련하였다.

나. 목적

본 안내서는 신용정보회사, 본인신용정보관리회사, 채권추심회사, 신용정보집중기관 및 신용정보제공·이용자(이하 '신용정보회사등')가 개인신용정보를 가명처리 또는 익명처리할 때 참고할 수 있는 사항을 안내하는 것을 목적으로 하며, 본 안내서에서 언급하지 않거나 안내서와는 다른 내용이더라도 가명·익명처리시 필요한 경우 관련법령을 준수하는 범위 내에서 신용정보회사등이 자체적으로 판단하여 활용할 수 있다. 본 안내서가 금융분야의 산업적 특성, 금융업권별 처리 정보의 특성 등을 고려하여 개인정보 자기결정권을 보장하고 금융산업 및 금융분야 정보산업의 발전에 기여할 수 있을 것으로 기대한다.

다. 적용범위

신용정보회사등이 개인신용정보를 가명처리·익명처리하거나 정보집합물 결합을 수행할 때 「신용정보법」, 「개인정보 보호법」 및 관계법령에 특별한 규정이 있는 경우를 제외하고는 본 안내서를 참고할 수 있다. 안내서의 내용과 법령이 상충될 경우 해당 법령에 따른다.

2. 용어정의

본 안내서에서 사용하는 용어의 뜻은 다음과 같다.

가. 개인신용정보

기업 및 법인에 관한 정보를 제외한 살아 있는 개인에 관한 신용정보로서 다음의 어느 하나에 해당하는 정보를 말한다(「신용정보법」 제2조제2호).

1) 해당 정보의 성명, 주민등록번호 및 영상 등을 통하여 특정 개인을 알아볼 수 있는 정보

2) 해당 정보만으로는 특정 개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 특정 개인을 알아볼 수 있는 정보

◎ **「신용정보법」 제2조(정의)** 1. "신용정보"란 금융거래 등 상거래에서 거래 상대방의 신용을 판단할 때 필요한 정보로서 다음 각 목의 정보를 말한다.
가. 특정 신용정보주체를 식별할 수 있는 정보(나목부터 마목까지의 어느 하나에 해당하는 정보와 결합되는 경우만 신용정보에 해당한다)
나. 신용정보주체의 거래내용을 판단할 수 있는 정보
다. 신용정보주체의 신용도를 판단할 수 있는 정보
라. 신용정보주체의 신용거래능력을 판단할 수 있는 정보
마. 가목부터 라목까지의 정보 외에 신용정보주체의 신용을 판단할 때 필요한 정보

나. 속성(attribute)

속성이란 데이터의 고유한 특성을 말하며, 다음과 같이 구분된다.

1) 식별자

주민등록번호, 이메일주소, 휴대전화번호 등과 같이 그 자체로 특정 개인을 직접 식별하는 용도로 사용하는 속성을 말한다.

2) 개인식별가능정보

연령, 성별, 거주지역, 국적 등과 같이 해당 정보만으로는 직접적으로 특정 개인을 식별할 수 없지만, 다른 속성과 결합하여 특정 개인의 신원을 전부 또는 일부를 드러낼 수 있는 속성을 말한다. 이러한 개인식별가능정보는 다른 속성과 결합할 경우 개인 식별 가능성이 높은지 낮은지에 따라 가명처리·익명처리의 수준 등을 달리할 수 있으며, 해당 속성이 개인 식별 가능성이 높은지 여부는 구체적인 사례에 따라 달리 판단될 수 있다.

< (예시) 속성의 분류 >

속성	가명·익명처리 대상 정보
식별자	성명, 상세주소, 전화번호, 생체인식정보, 전자우편주소, 사회관계망 서비스 주소, 주민등록번호, 여권번호, 운전면허번호, 외국인등록번호, 「정보통신망법」 제23조의3에 따른 본인확인기관이 특정 개인을 고유하게 식별할 수 있도록 부여한 정보, 특정 개인을 고유하게 식별하거나 동일한 신용정보주체를 구분하기 위하여 부여된 정보, 국내거소신고번호, 계좌번호, 신용카드번호, 건강보험증번호, 기기식별자, 자동차번호 등
개인식별가능 정보	성별, 나이, 주소, 우편번호, 직업(직업명 혹은 직업코드), 사건발생일자(사망, 승인, 수술, 퇴원, 방문 등), 위치(우편번호, 건물명, 지역 등), 인종, 출생국, 모국어, 가시적 소수인종집단 지위(visible minority status), 결혼 여부, 학력, 범죄경력, 종교, 의료 진단명, 보험 가입정보(보험 종류, 가입건수, 가입채널, 가입일, 보장금액 등), 신용대출 정보(대출건수, 계약일, 대출액, 상환액, 연체율 등), 납입보험료, 추정소득, 추정주택가격, 보유차량 정보, 핵심고객 여부, 내부 신용등급, CB신용점수 등

※ 동 예시는 식별자 또는 개인식별가능정보에 해당할 수 있는 속성을 예시한 것으로 실제로 개별 속성이 식별자 또는 개인식별가능정보에 해당하는지 여부는 개별 사례, 이용환경 등에 따라 상이할 수 있음

다. 식별(identification)

단독으로 또는 두 개 이상의 속성을 결합하는 등의 방법으로 개인을 알아볼 수 있도록 처리*하는 것을 말한다.

* "처리"란 신용정보의 수집(조사를 포함한다. 이하 같다), 생성, 연계, 연동, 기록, 저장, 보유, 가공, 편집, 검색, 출력, 정정(訂正), 복구, 이용, 결합, 제공, 공개, 파기(破棄), 그 밖에 이와 유사한 행위를 말한다.(동법 제2조제13호)

라. 정보집합물

정보를 체계적으로 관리하거나 처리할 목적으로 일정한 규칙에 따라 구성되거나 배열된 둘 이상의 정보들을 말한다(동법 제2조제15호 나목).

마. 결합키

결합의뢰기관이 정보집합물을 데이터전문기관에 제공하는 경우 하나의 정보집합물과 다른 정보집합물간에 둘 이상의 정보를 연계, 연동하기 위하여 사용되는 정보로, 해당 개인을 식별할 수 없으나 구별할 수 있는 정보를 말한다.

바. 가명처리

추가정보(예 : 가명정보와 기존 식별자를 연결하는 매핑테이블 등)를 사용하지 아니하고는 특정 개인인 신용정보주체를 알아볼 수 없도록 개인신용정보를 처리하는 것을 말하는데, 그 처리 결과가 ① 어떤 신용정보주체와 다른 신용정보주체가 구별되는 경우 ② 하나의 정보집합물에서나 서로 다른 둘 이상의 정보집합물 간에 어떤 신용정보주체에 관한 둘 이상의 정보가 연계되거나 연동되는 경우 ③ 위와 유사한 경우로서 대통령령으로 정한 경우의 어느 하나에 해당하는 경우로서 법령에 따라 그 추가정보를 분리하는 등 특정 개인인 신용정보주체를 알아볼 수 없도록 개인신용정보를 처리한 경우를 포함한다(동법 제2호제15호).

※ 개별 속성(식별자, 개인식별가능정보 등)에 대한 가명처리 수준은 가명처리 목적 및 가명정보의 이용환경과 가명정보 및 추가정보에 대한 보호조치 수준에 따라 달라질 수 있으며, 가명처리 단계에서 가명정보 이용환경(제3자 제공 여부, 외부 공개여부 등)과 기술적·관리적·물리적 보호조치 수립 여부 등을 고려하여 적절한 가명처리 수준을 결정해야 한다.

사. 가명정보

가명처리한 개인신용정보를 말한다(동법 제2조제16호).

아. 추가정보

특정 개인인 신용정보주체를 알아볼 수 없도록 개인신용정보를 가명처리 하는데 사용된 정보로서 가명정보를 원래의 상태로 복원하는데 사용할 수 있는 값을 말한다.

※ 예시 : 가명정보와 기존 식별자를 연결하는 매핑테이블, 가명정보 생성시 사용한 암호 알고리즘, 가명정보 생성시 사용한 솔트값 등

자. 익명처리

데이터 값 삭제, 가명처리, 총계처리, 범주화 등의 방법으로 개인신용정보의 전부 또는 일부를 삭제하거나 대체함으로써 더 이상 특정 개인인 신용정보주체를 알아볼 수 없도록 개인신용정보를 처리하는 것을 말한다(동법 제2조제17호).

차. 익명정보

개인신용정보를 더 이상 특정 개인인 신용정보주체를 알아볼 수 없도록 익명처리한 정보를 말한다.

카. 연결키

결합의뢰기관이 결합된 정보집합물을 이용하여 시계열 분석·연구 등을 수행할 수 있도록 결합키를 유일하게 대체하는 값을 말한다.

3. 개인정보, 가명정보, 익명정보

「신용정보법」은 가명처리와 익명처리 개념을 도입하였다. 개인정보에서 가명정보, 익명정보로 갈수록 식별가능성이 낮아지게 된다. 개인정보, 가명정보 및 익명정보를 개념 및 활용 가능한 범위를 기준으로 구분하면 다음과 같다.*

구분	개념	활용가능 범위
개인정보	특정 **개인에 관한 정보,** 개인을 **알아볼 수 있게 하는 정보**	**정보주체**로부터 **사전**에 **구체적**인 **동의**를 받은 범위 등의 내에서 활용 가능
가명정보	**추가정보**의 사용 없이는 **특정 개인을 알아볼 수 없게 조치**한 정보	다음 목적으로는 **동의 없이** 활용 가능 ❶ 통계작성(상업적 목적 포함) ❷ 연구(산업적 연구 포함) ❸ 공익적 기록보존 목적 등
익명정보	**더 이상 개인을 알아볼 수 없게 조치**한 정보	개인정보가 아니기 때문에 **제한 없이 자유롭게 활용**

※ 구체적인 개인정보, 가명정보, 익명정보의 예시는 아래의 [참고] 참조

* 금융위원회 보도참고자료, "「신용정보법」 개정으로 데이터를 가장 안전하게 잘 쓰는 나라를 만들겠습니다", 2019.11.28.

[참고] 가명정보와 익명정보 예시

□ 원본 정보집합물 정보

< (예시) 원본 정보 >

성명	전화번호	성별	생년월일	보험가입건수
신사임당	010-1234-5678	여	1974.10.1.	3
권율	02-2345-6789	남	1990.3.26.	2
유관순	010-3456-4321	여	1969.5.28.	1
이순신	010-4567-9876	남	1993.11.3.	2
선덕여왕	010-5678-9012	여	1971.1.2.	3
안중근	010-6789-0123	남	1988.7.16.	3
류성룡	010-7890-1234	남	1994.2.3.	2
이황	010-8901-2345	남	1982.6.28.	5
이이	010-9012-3456	남	1985.8.5.	2
…	…	…	…	…

○ (식별자) ‘성명’과 ‘전화번호’는 직접적으로 개인 식별 가능

○ (개인식별가능정보) ‘성별’과 ‘생년월일’은 다른 정보와 조합하여 개인을 식별할 가능성이 높고, ‘보험가입건수’는 통계 정보로 개인이 식별될 가능성이 낮음

❶ 가명정보

< (예시) 가명처리된 정보 >

ID	~~성명~~	~~전화번호~~	성별	출생년도	보험 가입건수
9A00F1155584BA5DDFFC4B6DDD7940431737C612651267FBD4716FE93C46F6BA	~~신사임당~~	~~010-1234-5678~~	여	1974	3
C2E6376B9035D7067C8B68F25FA34592F210D72E59B8E3F018C941B391AB1D99	~~권율~~	~~02-2345-6789~~	남	1990	2
DACE2CCC9F459387EAE890D8534955003F78B2B474C997CF2D990573D4C3344F	~~유관순~~	~~010-3456-4321~~	여	1969	1
27B339D75FF1DCED2C29A866BA5D61555D4C2E2C708F121AFABF34E5777AE498	~~이순신~~	~~010-4567-9876~~	남	1993	2
6CE926B166980F9C5F05F0B19A443E3494943BDACF2A657DFA1B2CF37C17B839	~~선덕여왕~~	~~010-5678-9012~~	여	1971	3
05CF80408DCC19A18228A365BD2DBBD4328BC36DC832F6E7365E536164A92B5A	~~안중근~~	~~010-6789-0123~~	남	1988	3
11834268AF3110DB64360198755400A49AF1A60A0BFE624DCE108B9E1185FA6C	~~류성룡~~	~~010-7890-1234~~	남	1994	2
725F8676075F7C0C5E6655EE84FF0EA2BEFD57D7F6C338083A961C211AAE952D	~~이황~~	~~010-8901-2345~~	남	1982	5
380A314D13F03BB6DBBAA0EAC76E26C1ED3A19A7AA74661162861D021FDEED7E	~~이이~~	~~010-9012-3456~~	남	1985	2
...	...	...	...	...	...

○ 성명, 전화번호, 성별, 생년월일을 조합하여 가명처리 기법 중 하나인 해시함수(SHA-256, 솔트값)를 적용

○ 식별자(성명, 전화번호)는 삭제하고 개인식별가능정보(성별, 생년월일)는 활용하되 개인 식별 가능성이 높은 성별, 생년월일 등은 일반화 처리* 가능

* 가명정보 이용자의 개인정보 보호수준과 가명정보의 재식별 가능성 등에 따라 가명처리 수준은 달라질 수 있음(위험도가 높을수록 가명처리 수준도 높아짐)

(예) 원본정보(1974.9.23.) → 출생년도만 남김(1974년) → 연령대로 범주화(40대)

❷ 익명정보

< (예시) 익명처리된 정보 >

~~성명~~	~~전화번호~~	성별	나이	보험 가입 건수	
~~권율~~	~~02-2345-6789~~	D	20대	2	} 동질집합 (k=3)
~~이순신~~	~~010-4567-9876~~	D	20대	2	
~~류성룡~~	~~010-7890-1234~~	D	20대	2	
~~안중근~~	~~010-6789-0123~~	D	30대	3	} 동질집합 (k=3)
~~이황~~	~~010-8901-2345~~	D	30대	5	
~~이이~~	~~010-9012-3456~~	D	30대	2	
~~신사임당~~	~~010-1234-5678~~	C	40대	3	} 동질집합 (k=3)
~~유관순~~	~~010-3456-4321~~	C	40대	1	
~~선덕여왕~~	~~010-5678-9012~~	C	40대	3	
		…	…	…	

○ 식별자(성명, 전화번호)는 삭제

○ 개인식별가능정보 중 다른 속성과 결합할 때 개인 식별 가능성이 높은 ‘성별’은 직접 알아볼 수 없도록 코드 형태로 변환(여성→C/남성→D)

○ 개인식별가능정보 중 다른 속성과 결합할 때 개인 식별 가능성이 높은 ‘생년월일’은 k-익명성*을 충족하기 위해 생일을 삭제하고, 나이를 연령대로 범주화

* k값은 익명정보 이용 목적, 환경 등에 따라 상이

※ 본 안내서 ‘Ⅲ. 2. 익명처리 방법’을 참고

○ ‘보험 가입 건수’는 분석 대상이 되는 속성이고 다른 속성과 결합할 때 개인 식별 가능성이 낮다고 판단되어 변환하지 않음

4. 금융분야 가명·익명처리 일반

가. 가명처리

1) 가명정보의 범위

가명정보는 가명처리한 개인신용정보를 말하며, 어떤 신용정보주체와 다른 신용정보주체가 '구별'되더라도 특정 신용정보주체를 식별할 수 없는 경우에는 가명정보로 볼 수 있다. 또한, 가명정보는 하나의 정보집합물 내에서 또는 서로 다른 둘 이상의 정보집합물 간에 어떤 신용정보주체에 관한 둘 이상의 정보가 연계되거나 연동되는 경우를 포함하며, 별도의 매핑테이블 등 추가정보가 존재하는 경우에도 가명처리한 정보는 가명정보로 볼 수 있다(동법 제2조제15호).

※ 구별 : 성질이나 종류가 차이가 나는 것을 의미하는 것으로 특정 속성이 다른 속성과 구분되는 것을 의미함

※ 식별 : 분별하여 알아보는 것을 의미하는 것으로 속성을 통해 개인을 알아볼 수 있는 것을 의미함

그러나 위 경우에도 추가정보를 사용하지 아니하고는 특정 개인인 신용정보주체를 알아볼 수 없어야 하며, 법령이 정한 가명처리에 관한 행위규칙을 준수해야 한다. 개별 속성에 대한 가명처리 수준은 가명처리 목적, 제3자 제공 여부, 외부 공개 여부 등과 가명정보 및 추가정보에 대한 기술적·관리적·물리적 보호조치 수준 등에 따라 달라질 수 있다.

2) 가명정보의 활용

가명정보는 통계작성(상업적 목적을 포함), 연구(산업적 연구를 포함), 공익적 기록보존 등을 위하여 가명정보를 제공하는 경우에는 개인인 신용정보주체의 동의 없이 가명정보를 활용할 수 있다(동법 제32조제6항 제9호의2). 이 경우 통계작성에는 시장조사 등 상업적 목적으로 수행

하는 통계작성을 포함하며, 연구에는 대학, 연구소 등 연구기관 뿐 아니라 기업 등이 수행하는 산업적 연구를 포함한다. 다만, 특정 개인을 식별할 수 있는 형태의 통계작성, 연구, 공익적 기록 보존 등의 행위는 모두 허용되지 않는다.

○ **(통계작성)** 집단적 현상이나 수집된 자료의 내용에 관한 수량적인 정보를 작성하는 행위

> **◎ 예시**
> ▶ 금융기관 소액대출 심사의 신용 보조지표로 활용하기 위하여 고객·지역별 신용카드 결제 데이터, 아파트 관리비, 부동산 시세 등에 대한 통계를 작성하는 경우
> ▶ 지자체 쓰레기 수거량을 예측하기 위하여 신용카드 결제건수·이용금액, 가맹점 업종·지역, 고객 거주·직장 지역, 거주 지역별 온·오프라인 구매물품, 배달 음식 매출액·건수 등에 관한 통계를 작성하는 경우

○ **(연구)** 기술 개발, 실증, 기초연구, 응용연구, 민간투자연구 등 과학적 방법을 적용하는 연구를 의미

- 자연과학적 연구뿐만 아니라 과학적 방법을 적용하는 역사적 연구, 공중보건 분야에서 공익을 위해 시행되는 연구 등은 물론, 새로운 기술·제품·서비스의 개발, 시장조사 등 산업적 목적의 연구도 포함

> **◎ 예시**
> ▶ 보험사기 자동 탐지시스템 개발을 위하여 과거 10년간의 보험사기 사례에 대한 보험금 청구금액, 청구시점과 방법, 유사청구 반복 여부 등을 분석하여 보험사기의 징후를 발견하기 위한 연구를 하는 경우

○ **(공익적 기록보존)** 공공의 이익을 위하여 지속적으로 열람할 가치가 있는 기록정보를 보존하는 것을 의미

- 공공기관이 처리하는 경우에만 공익적 목적이 인정되는 것은 아니며, 민간기업, 단체 등이 일반적인 공익을 위하여 기록을 보존하는 경우도 공익적 기록 보존 목적이 인정됨

◎ **예시**

▶ 연구소가 현대사 연구 과정에서 수집한 개인정보 중에서 사료가치가 있는 인물정보를 기록하여 보관하는 경우

3) 가명처리 관련 의무 및 처벌 규정

신용정보회사등은 가명처리에 사용한 추가정보를 기술적 · 관리적 · 물리적 보호조치를 통해 추가 정보에 대한 접근을 통제하는 방법으로 분리하여 보관하거나 삭제하여야 하며(동법 제40조의2제1항), 가명처리한 개인신용정보에 대하여 제3자의 불법적인 접근, 입력된 정보의 변경·훼손 및 파괴, 그 밖의 위험으로부터 가명정보를 보호하기 위하여 내부관리계획을 수립하고 접속기록을 보관하는 등 기술적·관리적·물리적 보안대책을 수립·시행하여야 한다(동법 제40조의2제2항).

또한 신용정보회사등은 영리 또는 부정한 목적으로 특정 개인을 알아볼 수 있게 가명정보를 처리하여서는 아니 되며(동법 제42조의2제6항), 신용정보회사등이 영리 또는 부정한 목적으로 특정 개인을 알아볼 수 있게 가명정보를 처리한 경우, 금융위원회는 관련 매출액이 아닌 '전체 매출액'의 100분의 3 이하에 해당하는 금액을 과징금으로 부과할 수 있다(동법 제42조의2제1항 제1호의4).

그리고 영리 또는 부정한 목적으로 특정 개인을 알아볼 수 있게 가명정보를 처리한 자는 5년 이하의 징역 또는 5천만원 이하의 벌금에 처해질 수 있다(동법 제50조제2항 제7호의2).

* 동법 제40조의2 제1항 내지 제2항의 사항은 동법 시행령 제34조의5제1항 내지 제3항에 정하고 있으며, 「신용정보업 감독규정」 제43조의7 및 [별표 8]에 금융위원회는 그 세부사항을 규정(본 안내서 'Ⅱ. 4. 가명정보 및 추가정보에 관한 보호조치 기준' 참조)

한편, 가명처리에 사용한 추가정보를 분리하여 보관하거나 삭제하지 아니한 자, 가명처리한 개인신용정보에 대하여 기술적·관리적·물리적 보안대책을 수립·시행하지 아니한 자, 가명정보를 이용하는 과정에서 특정 개인을 알아볼 수 있게 된 경우 즉시 그 가명정보를 회수하여 처리를 중지하거나 즉시 삭제하지 아니하는 자는 3천만원 이하의 과태료에 처해질 수 있다(동법 제52조제3항 제16호 내지 제18호).

나. 익명처리

1) 익명정보의 활용 범위

익명정보는 특정 개인인 신용정보주체를 알아볼 수 없도록 개인신용정보를 처리한 것으로, 개인을 알아볼 수 없는 정보임을 전제(동법 제2조제17호)로, 별도의 제한 없이 사용할 수 있다.

2) 익명정보의 적정성 평가

신용정보회사등은 개인신용정보에 대한 익명처리가 적정하게 이루어졌는지 여부에 대하여 금융위원회에 그 심사를 요청할 수 있다(동법 제40조의2제3항). 금융위원회가 위 요청에 따라 심사하여 적정하게 익명처리가 이루어졌다고 인정한 경우 더 이상 해당 개인인 신용정보주체를 알아볼 수 없는 정보로 추정한다(동법 제40조의2제4항). 금융위원회는 법 제40조의2제3항의 익명처리의 적정성 심사 및 법 제40조의2제4항의 익명처리의 적정성 인정업무를 데이터전문기관에 위탁한다(동법 시행령 제37조제5항).

다. 가명처리·익명처리의 조치 기록 보존 의무

신용정보회사등은 개인신용정보를 가명처리한 경우에는 가명처리한 날짜, 가명처리한 정보의 항목, 가명처리한 사유와 근거를, 개인신용정보를 익명처리한 경우에는 익명처리한 날짜, 익명처리한 정보의 항목, 익명처리한 사유와 근거를 3년간 보존하여야 한다(동법 제40조의2제8항).

위 보존의무를 위반하여 개인신용정보를 가명처리하거나 익명처리한 기록을 보존하지 아니한 자에게는 1천만원 이하의 과태료를 부과한다(동법 제52조제5항 제11호의3).

◎ 데이터전문기관의 업무

▶ 「신용정보법」 제26조의4 및 동법 시행령 제22조의4는 데이터전문기관의 업무를 다음과 같이 명시하고 있다.

「신용정보법」 제26조의4(데이터전문기관) ② 데이터전문기관은 다음 각 호의 업무를 수행한다.

1. 신용정보회사등이 보유하는 정보집합물과 제3자가 보유하는 정보집합물 간의 결합 및 전달
2. 신용정보회사등의 익명처리에 대한 적정성 평가
3. 제1호 및 제2호와 유사한 업무로서 대통령령으로 정하는 업무

「신용정보법」 시행령 제22조의4(데이터전문기관) ① 법 제26조의4제2항제3호에서 "대통령령으로 정하는 업무"란 다음 각 호의 업무를 말한다.

1. 정보집합물간의 결합과 가명처리 또는 익명처리에 대한 조사·연구 및 이와 유사한 업무
2. 정보집합물간의 결합과 가명처리 또는 익명처리의 표준화에 관한 사항
3. 데이터전문기관간 업무 표준화 등에 대한 상호 협력에 관한 사항
4. 그 밖에 이와 유사한 업무로서 금융위원회가 정하여 고시하는 업무

Ⅱ. 가명처리

1. 개요

식별자는 원칙적으로 삭제하여야 하며, 정보집합물 결합 등 데이터 이용목적 상 필요한 경우 안전한 방식으로 대체값을 생성하여 식별자를 대체하여야 한다.

개인식별가능정보는 금융분야에서 처리하는 개인(신용)정보 및 이용환경의 특성에 따라 개인식별가능정보 간의 조합, 외부에 공개된 정보와의 결합, 특이치(outlier)* 등으로 인하여 개인 식별 가능성이 높은 경우 일반화, 범주화 등의 추가적인 조치를 통해 재식별 위험을 낮춰야 한다. 또한 개별 속성에 대한 가명처리 수준은 가명처리 목적, 가명정보 이용환경과 가명정보 및 추가정보에 대한 보호조치 수준에 따라 달라질 수 있다. 예를 들어, 내부 연구목적으로 가명정보를 활용하려는 경우 생년월일 정보를 출생년도(1974년)로 이미 처리했더라도, 동일한 가명정보를 제3자에게 제공하는 경우에는 가명처리 수준을 연령대(40대)로 높일 수 있다. 개인식별가능정보는 개인 식별의 위험이 높지 않다면 원칙적으로 별도의 조치 없이 사용할 수 있다. 다만, 이용 상황에 따라 개인 식별 가능성이 높아진 경우에는 추가적인 조치를 통해 재식별 위험을 낮춰야 한다.

* 관측된 데이터의 범위에서 많이 벗어난 아주 작은 값이나 아주 큰 값

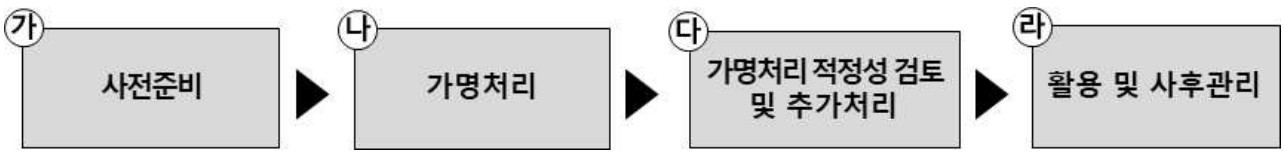
◎ 특이치 예시

- ▶ 신용정보주체의 연령이 20세에서 75세 사이에 대부분 분포하는 데이터셋에서 특정 신용정보주체의 연령만 110세인 경우
- ▶ 신용정보주체의 대출금액이 5백만원에서 10억원 사이에 대부분 분포하는 데이터셋에서 특정 신용정보주체의 대출금액이 80억원인 경우

2. 가명처리 절차

신용정보회사등은 다음의 절차 예시 등을 참고하여 가명정보의 위험도를 검토하고 그에 따른 적절한 가명처리를 수행하여야 한다.

< (예시) 가명처리 단계별 절차 >



가. 사전준비

가명처리 목적을 명확히 정의하여 그에 따른 가명처리 대상 데이터를 추출하고, 가명정보 처리 및 활용에 대한 내부 체계를 구축하여야 한다. 신용정보회사등이 가명정보를 제3자에게 제공할 경우, 사후 책임문제를 명확하게 하기 위하여 재식별 금지, 정보유출시 손해배상 등을 반영한 계약서 등을 작성할 필요가 있다.

○ **(가명처리 목적 명확화)** 가명정보의 활용 목적은 「신용정보법」에서 허용하는 목적 내에서 최대한 구체화

※ 통계작성(상업적 목적 포함), 연구(산업적 연구 포함), 공익적 기록보존 목적 등

○ **(처리대상 추출)** 가명처리 목적을 달성하기 위해 반드시 필요한 최소한의 항목으로 가명처리 대상 정보집합물을 추출

○ **(가명정보 처리 및 활용 체계 구축)** 가명정보 활용에 대한 관리방안을 수립하고 가명정보 및 추가정보에 대한 접근관리 체계를 구축*

* 가명정보 및 추가정보에 대한 접근통제 등 관리방안 수립 등('4. 가명처리에 관한 행위규칙' 참고)

나. 가명처리

가명처리 및 가명정보 이용 환경, 가명처리 대상 데이터의 특성 등을 고려하여 위험도를 측정하고 가명처리 수준을 결정한 후 가명처리를 수행한다.

○ **(위험도 측정)** 가명처리 목적, 처리·이용 환경 및 가명처리 대상 데이터의 특성 등에 따른 위험도 분석

< (예시) 위험도 측정시 고려사항 >

고려사항	세부 내용
가명처리 목적	- 통계작성(상업적 목적 포함), 연구(산업적 연구 포함), 공익적 기록보존 목적 여부 등 범위내에서 세부적인 목적
가명정보 활용 주체	- 내부 활용/내부 결합/외부 제공/외부 결합/외부 공개 여부 등
가명처리·이용 환경	- 처리 환경 및 이용(분석) 환경의 내부통제 수준, 재식별 의도 또는 능력 등 ※ 'Ⅱ. 3. 다. 가명정보의 재식별 위험도 측정시 고려사항' 참고
가명처리 대상 데이터의 특징 분석	- 가명처리 대상 데이터의 특성 분석 - 데이터 속성(컬럼)을 식별자, 개인식별가능정보 등으로 분류 ※ 식별자, 개인식별가능정보 예시는 'Ⅰ. 2. 나. 속성' 참고

○ **(가명처리 수준 결정)** 위험도를 고려하여 적절한 가명처리 방법·수준을 결정하고 가명정보 및 추가정보의 보유기간을 정의*

* 가명정보의 이용목적, 가명처리의 기술적 특성, 정보의 속성, 추가정보에 대한 기술적·관리적·물리적 보호조치 수준, 가명정보의 재식별시 신용정보주체에 미치는 영향, 가명정보의 재식별 가능성, 가명정보의 이용목적 및 그 목적 달성에 필요한 최소기간 등을 고려

○ **(가명처리)** 식별자의 삭제 또는 대체*, 재식별 위험도가 높은 개인식별가능정보에 대한 가명처리** 등을 수행

* 대체값 생성시 랜덤값 생성, 해시값 생성, 암호화 등 안전한 방식 활용 필요 ('Ⅱ. 3. 가. 식별자의 대체값 생성 방법' 및 'Ⅱ. 3. 나. 속성별 가명처리 방법' 참고)

** 일반화, 범주화, 상·하단 코딩, 레코드 삭제 등의 기법 활용('붙임 1. 가명·익명처리 기법' 참고)

다. 가명처리 적정성 검토 및 추가처리

앞의 '나 단계(가명처리)'에서 가명처리 수준이 적절히 정의되었고 이에 따라 가명처리가 제대로 되었는지 여부를 확인하고, 재식별 가능성 등을 검토하여 필요시 추가로 가명처리를 수행한다.

○ **(적정성 검토)** 내부자 또는 필요시 외부 전문가를 활용하여 가명정보의 개인 식별 가능성(식별자 존재 여부, 가명처리 수준의 적절성 등)을 검토

◎ 예시

▶ 연구 목적으로 가명처리한 데이터를 대상으로 내부 개인정보보호 책임자 및 외부의 법률 전문가 1명, 가명·익명처리 전문가 1명을 포함한 평가회의를 개최하여 가명처리의 적정성을 검토

※ 본 절차는 필수 사항은 아니며, 필요에 따라 신용정보회사등이 자체 절차를 수립하여 이행할 수 있음

라. 활용 및 사후관리

가명정보를 이용·제공·결합한 후 가명정보의 파기 등 가명정보 활용에 대한 행위규칙 등을 준수한다.

※ 'Ⅱ. 4. 가명처리에 관한 행위규칙' 및 'Ⅱ. 5. 가명정보 및 추가정보에 관한 보호조치 기준' 참고

< (예시) 가명처리 세부 절차 >

번호	단계	내용
가. 사전 준비	사전 준비	- 가명처리 목적 정의 - 가명처리 대상 정보집합물 추출 - 가명정보 및 추가정보에 대한 접근통제 등 관리방안 수립 등
나. 가명 처리	위험도 측정	- 가명처리 목적, 처리·이용환경(내부통제 수준, 재식별 의도 및 능력 등), 이용 주체(내부 활용/내부 결합/외부 제공/외부 결합/외부 공개 여부) 등에 따른 위험도(Risk) 분석 - 가명처리 대상 정보집합물의 특성 분석 - 데이터 속성(컬럼)을 식별자, 개인식별가능정보 등으로 분류 ※ 식별자, 개인식별가능정보 예시는 'Ⅰ. 2. 나. 속성' 참고
	가명처리 수준 결정	- 가명처리 방법 및 수준 결정 - 가명정보 및 추가정보의 보유기간 정의 ※ 가명정보의 이용목적, 가명처리의 기술적 특성, 정보의 속성, 추가정보에 대한 기술적·관리적·물리적 보호조치 수준, 가명정보의 재식별시 신용정보주체에 미치는 영향, 가명정보의 재식별 가능성, 가명정보의 이용목적 및 그 목적 달성에 필요한 최소기간 등을 고려
	가명처리	- 식별자에 대하여 삭제 또는 대체 - 대체값 생성시 안전한 방식 활용 필요 : 랜덤값 생성, 해시값 생성, 암호화 등 ※ 대체값 생성 알고리즘, 매핑테이블, 암호키 등 추가정보는 삭제 또는 분리보관 - 이용·제공 상황에 따라 재식별 리스크가 높다고 판단되는 경우, 개인식별가능정보에 대한 추가적인 가명처리* * 일반화, 범주화, 상·하단 코딩, 레코드 삭제 등의 기법 활용 (본 안내서 '붙임 1' 참고)
다. 적정성 검토	(필요시) 가명처리 적정성 검토	- 가명정보의 개인 식별 가능성 검토(식별자 존재 여부, 가명처리 수준의 적절성 등) - 내부자 검토 또는 (필요 시) 외부 전문가 활용 ※ 본 절차는 필수 사항은 아니며 필요에 따라 내부 절차를 수립하여 이행할 수 있음

번호	단계	내용
라. 활용 및 사후 관리	가명정보 이용·제공·결합 및 사후관리	- 통계작성, 연구, 공익적 기록보존 등의 목적으로는 신용정보주체의 동의 없이 가명정보 이용 또는 제공 가능 - 가명정보의 제3자 제공시 재식별 시도 금지, 책임 범위, 보호조치, 목적외 사용금지, 재제공 금지 등에 대한 사항을 계약 등을 통해 명시 필요 - 신용정보회사등의 경우 제3자와의 정보집합물 결합은 금융위원회가 지정한 데이터전문기관을 통해서만 가능 - 가명처리시 가명처리한 날짜, 정보의 항목, 사유와 근거를 기록하고 3년간 보존 - 「신용정보법」 제40조의2(가명처리·익명처리에 관한 행위규칙)에 따라 추가정보는 분리하여 보관하거나 삭제하고 가명정보를 안전하게 보호하기 위하여 내부관리계획을 수립하고 접속기록을 보관하는 등 기술적·관리적·물리적 보안대책을 수립·시행 - 가명정보를 이용하는 과정에서 특정 개인을 알아볼 수 있게 된 경우 즉시 그 가명정보를 회수하여 처리를 중지하고, 특정 개인을 알아볼 수 있게 된 정보는 즉시 삭제
	가명정보 삭제	- 가명처리 계획 수립시 정한 가명정보 보유 기간이 경과한 경우 삭제 조치 - 추가정보에 대해서도 반드시 필요하지 않은 경우 삭제 조치

3. 가명처리 방법

가. 식별자의 대체값 생성 방법

1) 원칙

식별자의 대체값(이하 ‘가명(pseudonym)’)은 일반적으로 랜덤값 생성, 해시값 생성, 암호화 기법 등을 활용할 수 있으며, 그 외에 이와 동일한 수준의 안전성을 확보할 수 있는 다른 방식(토큰화 등)을 활용할 수 있다.

가명 생성시 사용된 추가정보(매핑테이블, 암호키, 암호 알고리즘 등)*는 분리하여 보관하거나 삭제하는 등 「신용정보법」 제40조의2(가명처리·익명처리에 관한 행위규칙)에 따라 안전하게 관리되어야 한다.

* 가명과 원본의 식별자를 연결하는 매핑테이블을 생성할 수 있으며, 이 경우 매핑테이블이 추가정보에 해당

◎ 가명 생성시 입력정보로 식별자에 해당하는 CI(Connecting Information)값을 사용하는 경우, 전체 CI값을 일방향 해시함수의 입력정보로 사용할 수 없고 재식별 위험이 없는 범위 내에서 CI값의 일부를 사용하여야 한다.

2) 랜덤값 생성

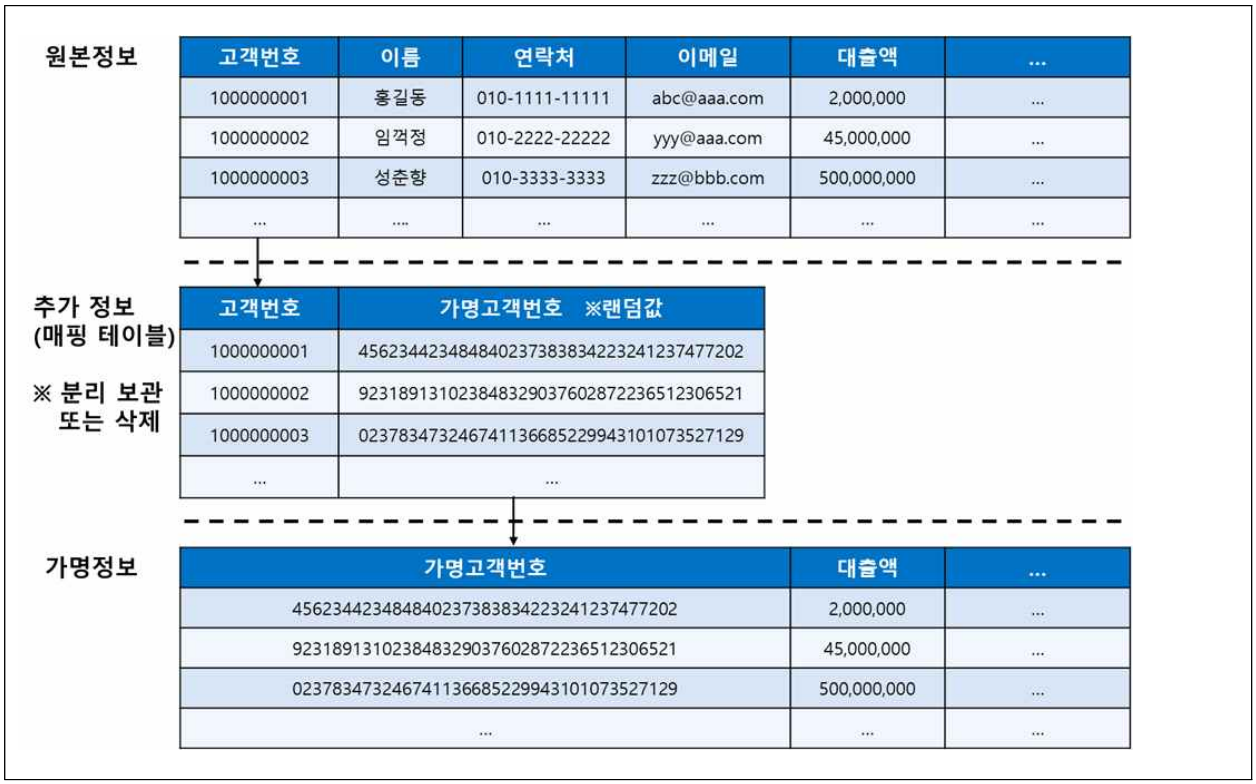
랜덤값 생성은 식별자에 대하여 독립적인 가명을 생성하는 방법으로, 랜덤값을 생성하여 원본값을 대체하는 방식을 말한다.

랜덤값 생성시 안전한 난수발생기(Random Number Generator, RNG)를 사용하여야 하며 난수 생성규칙이 노출되거나 중복이 발생하지 않도록 주의하여야 한다.

◎ 추가정보 예시

▶ 가명(랜덤값)과 원본 식별자를 연결하는 매핑 테이블을 생성할 수 있으며, 이 경우 매핑 테이블이 추가정보에 해당

< (예시) 랜덤값 생성을 통한 가명처리 예시 >



3) 해시값 생성

해시값 생성은 암호기술을 사용하여 식별속성으로부터 파생된 가명을 생성하는 방식으로서 단방향(one-way) 및 충돌 방지(collision-resistance) 특성을 가진 해시함수를 사용하여 단일 식별자 또는 다수 식별자들을 해시값으로 대체하는 방식을 말한다.

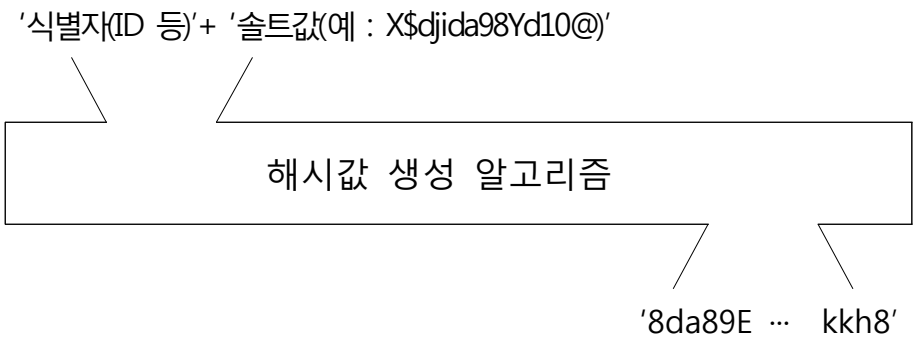
해시값에 대한 무작위 대입 공격*, 레인보우 테이블 공격** 등에 안전할 수 있도록 솔트값 또는 키값(key)을 추가하여 해시하여야 한다.

* 무작위 대입 공격(brute force attack) : 경우의 수를 무작위로 대입하여 원본값을 알아내는 공격

** 레인보우 테이블 공격(rainbow table attack) : 해시 함수를 사용하여 변환가능한 해시값을 미리 저장해 놓은 표를 통해 원본값을 알아내는 공격

해시값 생성시 사용되는 솔트값 또는 키값은 쉽게 유추할 수 없도록 복잡하게 구성하여야 하며 비인가자에게 노출되지 않도록 안전하게 관리하여야 한다.

< (예시) 해시 생성 예시 >



식별자에 단순히 솔트값 또는 키값만 추가하여 해시를 생성하기 보다는 자체적으로 해시값 생성규칙을 마련하여 해시값의 안전성을 높일 것을 권고한다.

◎ **해시값의 안전성을 높이는 해시값 생성규칙 예시**(H()를 해시함수로 가정)
▶ 예시 1 : H (솔트값1 + 식별자 + 솔트값2)
▶ 예시 2 : H (H (입력값 + 솔트값) + 솔트값) 등

해시값을 생성할 때 사용하는 해시 알고리즘은 SHA-2 이상의 안전성이 검증된 해시 알고리즘을 사용하여야 한다.

◎ **추가정보 예시**

▶ 필요시 가명(해시값)과 원본 식별자와의 매핑테이블을 생성·보관할 수 있으며 이때 사용한 솔트값 또는 키값, 해시값 생성규칙, 매핑 테이블 등이 추가정보에 해당

< 대표적인 권고 해시 알고리즘 >

종류		출력값 길이(비트)	보안강도(비트)	참조 규격
SHA-2	224	224	112	NIST FIPS 180-4
	256	256	128	
	386	384	192	
	512	512	256	
SHA-3	224	224	112	NIST FIPS 202
	256	256	128	
	386	384	192	
	512	512	256	

※ 출처 :「금융부문 암호기술 활용가이드」(금융보안원, 2019.1.)

4) 암호화

가명 생성시 식별자를 암호화하는 방식을 사용할 수 있다. 이 때, SEED, AES, ARIA 등 안전한 암호 알고리즘 사용하여야 한다. 특히, 암호 알고리즘은 컴퓨팅 파워 증가, 기술의 발전 등에 따라 안전성에 변화가 발생할 수 있으므로 가명 생성 시점에 안전하다고 평가 받는 암호 알고리즘을 확인하여 적용하여야 한다.

< 대표적인 대칭키 블록암호 알고리즘 >

종류	입출력 길이 (비트)	출력값 길이 (비트)	보안강도 (비트)	참조 규격
SEED	128	128	128	TTA TTAS.KO-12.0004/R1
AES	128	128	128	NIST FIPS 197
		192	192	
		192	192	

※ 출처 :「금융부문 암호기술 활용가이드」(금융보안원, 2019.1.)

암호화에 사용된 키가 유출될 경우 유출된 키를 통해 암호문을 복호화할 수 있으므로, 암호키 생성·배포·사용·정지·갱신·폐기 등의 암호키 관리 절차 수립, 암호키의 분리 보관, 암호키에 대한 접근통제 조치 등 안전한 암호키 관리방안이 수립·이행되어야 한다.

◎ 추가정보 예시

▶ 암호화 방식에서는 암호키, 암호 알고리즘 등이 추가정보에 해당

나. 속성별 가명처리 방법

1) 식별자 조치사항

신용정보회사등은 가명처리할 때 정보집합물 내의 식별자는 삭제하거나 가명으로 대체하여야 한다. 금융분야에서 사용되는 주요 식별자의 예시는 다음의 표와 같다. 아래 예시 외에도 데이터 특성 및 이용환경 등에 따라 특정 신용정보주체를 식별할 수 있는 정보가 존재한다면 식별자에 해당할 수 있다.

< (예시) 금융분야 식별자 >

No	식별자 예시	설명	비고
1	성명		이름
2	상세주소		상세주소
3	전화번호		휴대폰번호, 집전화번호 등
4	바이오인식정보		지문, 홍채, 안면인식 등
5	전자우편주소		이메일 주소
6	사회관계망서비스 주소		SNS 주소
7	주민등록번호		개인식별번호

No	식별자 예시	설명	비고
8	여권번호		개인식별번호
9	운전면허번호		개인식별번호
10	외국인등록번호		개인식별번호
11	정보통신망법 제23조의3에 따른 본인확인기관이 특정 개인을 고유하게 식별할 수 있도록 부여한 정보		CI, DI
12	특정 개인을 고유하게 식별하거나 동일한 신용정보주체를 구분하기 위하여 부여된 정보		회원번호, 고객번호, 아이디, 멤버십번호 등
13	국내거소신고번호		-
14	계좌번호	개인과 유일하게 연결되는 정보로서 식별자에 해당	-
15	신용카드번호	개인과 유일하게 연결되는 정보로서 식별자에 해당	-
16	건강보험증번호	개인에게 유일하게 부여된 번호로서 식별자에 해당	-
17	기기식별자	모바일서비스 등을 통해 기기식별자를 수집하는 경우 회원정보와 결합되어 개인식별이 가능함	-
18	자동차번호	개인이 소유한 자동차번호의 경우 해당 정보를 바탕으로 개인을 식별할 가능성이 높음	-
19	그 밖에 특정 개인을 고유하게 식별할 수 있는 정보	데이터 특성, 이용 환경 등을 고려하여 특정 개인을 고유하게 식별할 수 있는 정보의 경우 식별자에 해당	사진, 영상 등

2) 개인식별가능정보 조치사항

개인식별가능정보 중 다른 정보와 결합할 경우 개인을 식별할 가능성이 높은 경우, 이용·제공 목적상 반드시 필요하지 않은 개인식별가능정보는 삭제하고 나머지 개인식별가능정보에 대해서는 적절한 수준의 추가 조치를 적용해야 한다. 금융분야에서 사용되는 주요 개인식별가능정보 및 조치사항 예시는 다음의 표와 같다.

< (예시) 금융분야 주요 개인식별가능정보 및 조치사항 >

No	개인식별가능정보	조치 사항 예시	비고
1	성별	- 이용 목적상 필요하다면 별도 조치 없이 사용 가능	-
2	나이	- 필요시 상황에 따라 5세, 10세 간격 등으로 범주화 - 특정 나이 이상 또는 이하의 경우 단일 범주로 집계(상·하단 코딩)	범주화, 상·하단코딩 등
3	주소	- 세부주소의 경우 식별자에 해당하므로, 필요시 시·군·구 단위 등으로 범주화 - 특히 도서산간 등 일부 지역의 경우 읍·면·동 단위의 거주자가 매우 적을 수 있으므로 필요시 범주화 등 조치 필요 - 우편번호에 대해서도 동일한 기준 적용	범주화 등
4	직업	- 국회의원, 연예인, 운동선수 등 일부 직업의 경우 개인 식별 가능성이 높아지므로, 필요시 직업 분류에 명시적으로 드러나지 않도록 조치	일반화, 범주화 등
5	국적	- 특정 집단 내에서 대다수가 동일 국적자인 경우, 그 외 국적자에 대한 개인 식별 가능성이 높아지므로 필요시 범주화 등 조치 필요	범주화 등
6	기념일	- 결혼기념일 등 일부 기념일의 경우 개인 식별 가능성이 높아지므로 필요시 범주화 등 조치	범주화 등

No	개인식별가능정보	조치 사항 예시	비고
7	기혼여부	- 이용 목적상 필요하다면 별도 조치 없이 사용 가능	범주화 등
8	거래지점	- 거래지점의 경우 거래자의 주요 활동지를 한정지을 수 있으므로 필요시 범주화 등 조치 필요 - 필요시 거래지점명/거래지점코드 대신 거래지점이 위치한 구 단위, 동 단위 주소 등으로 대체	범주화 등
9	기타	- 데이터 특성 상 다른 정보와 결합하여 개인을 식별할 가능성이 높은 속성이 존재한다면 개인 식별 가능성이 높은 개인식별가능정보로 지정 - 개인식별가능정보가 데이터 특성, 이용 상황, 제3자 제공 여부 등에 따라 재식별 위험이 높다고 판단되는 경우 추가 조치 적용	일반화, 범주화, 잡음추가, 삭제, 상하단코딩 등

위의 예시에도 불구하고 데이터 특성 및 이용환경 등에 따라 추가 조치 필요 여부, 조치 방법 및 수준 등이 상이할 수 있으므로, 재식별 위험에 근거하여 적절한 조치를 취하여야 한다. 특히, 정보집합물을 외부에 제공하거나 결합하는 경우에는 재식별 위험이 높아질 수 있으므로 추가적인 조치의 적용을 고려할 필요가 있다.

다. 가명정보의 재식별 위험도 측정시 고려사항

가명정보의 재식별 위험도는 가명처리 수준을 결정하기 위한 주요한 요소이다. 가명정보 보호수준이 높은 이용기관에는 낮은 수준으로 가명처리 된 가명정보를 제공하여 활용성을 높이고, 보호수준이 낮은 이용기관에는 높은 수준으로 가명처리된 가명정보를 제공하여 개인신용정보가 재식별 될 수 있는 가능성을 줄여야 한다.

신용정보회사등이 이용기관의 가명정보 보호수준을 판단하기 위해서는 해당 이용기관의 재식별 의도와 능력, 가명정보 보호능력, 업무수행 신뢰도 등 다양한 측면에서 평가가 필요하며 이를 종합적으로 고려하여 가명처리 수준을 결정해야 한다.

1) 재식별 의도 및 능력 분석

가명정보 이용기관의 재식별 의도 및 능력에 대해 검토하고, 재식별 의도와 능력이 높게 평가될 경우 가명처리 수준을 높일 필요가 있다.

○ **(재식별 의도)** 가명정보 이용자가 가명정보를 재식별하여 경제적·비경제적 이득을 취할 수 있거나, 목적에 부합하지 않는 범위로 가명정보를 활용할 여지가 있는지 등을 검토

○ **(재식별 능력)** 가명정보 이용자가 재식별을 시도할 수 있는 전문지식이나 가명정보와 연계 가능한 데이터를 보유하고 있는지 등을 검토

2) 가명정보 보호수준 및 신뢰도 분석

가명정보도 개인(신용)정보로 볼 수 있으므로, 가명정보 이용기관의 가명정보 보호수준 및 업무수행 신뢰도에 대해 검토하고, 가명정보 보호능력과 업무수행 신뢰도가 낮게 평가될 경우 가명처리 수준을 높일 필요가 있다.

○ **(가명정보 보호수준)** 가명정보 이용기관이 가명정보를 보호하기 위한 가명정보 관리계획을 수립·운영하고 기술적·관리적·물리적 보호조치를 마련하였는지, 개인정보 보호 관련 인증 유무 등을 검토

※「신용정보업 감독규정」[별표 8] 가명정보에 관한 보호조치 기준 외에도, 동 규정의 [별표 3] 기술적·관리적·물리적 보안대책 마련 기준,「개인정보보호법」 및 동 법률에 따른「개인정보의 안전성 확보조치 기준」등 관계 법령의 개인(신용)정보 보호를 위한 조치 기준을 준수하는지 여부에 대하여 판단

○ **(업무수행 신뢰도)** 가명정보를 활용하면서 위법을 저지른 적은 없는지, 가명정보를 다른 기관에게 무단으로 제공할 가능성은 없는지 등을 검토

3) 분석 결과 해석

재식별 의도나 능력이 높다고 해서 무조건적으로 가명처리 수준을 높일 필요는 없다. 예를 들어 재식별 능력이 높다고 하더라도 가명정보 보호능력과 업무수행 신뢰도가 매우 높다면 가명처리 수준을 낮출 수 있을 것이다.

반대로 가명정보 보호능력과 업무수행 신뢰도가 높다고 해서 반드시 가명처리 수준을 낮춰서는 안 된다. 높은 보호능력을 가지고 있더라도 재식별 의도가 높아 보이는 등의 경우 가명처리 수준을 높여야 할 것이다.

[가명처리 예시 ①] 내부 활용

▷ 처리 목적 : A카드회사는 2021년 상반기 출시 예정인 중금리 대출 상품을 준비하면서 대출심사전략을 마련하고자 2019년 카드이용고객 중 중금리 상품 이용 예상 고객을 추출하여 관련 분석·연구를 추진

▷ 가명정보 활용구간 : 내부

▷ 보호조치·접근통제 수준 : 가명정보 및 추가정보에 대해 내부관리계획에 따라 개인신용정보에 준하는 기술적·관리적·물리적 보안 대책을 적용하고 내부통제 하에 활용

가. 보유 정보 샘플

ID	성명	카드 번호	전화 번호	성별	생년 월일	주소	근무처	연봉 (만원)	내부 신용 등급	연체 잔고 (만원)	결제 기관
19342	홍길동	3779 4593 3043 3921 3943	010 -3355 -0934	남	1972. 9.9.	서울시 강남구 역삼동 332-1	OO 자동차	4,500	5	35	국민 은행
19354	김철수	4832 2332 2344 4399	02 -531 -9834	남	1980. 4.16.	서울시 마포구 송내동 334-1	OO 은행	6,000	2	0	새마을 금고
19445	전지연	4523 3234 9843 0394	010 -9290 -3344	여	1979. 5.23.	경기도 광주시 송정동 786-1	OO 공사	5,500	9	125	우리 은행
20221	박식별	4932 3453 5943 4321	010 -2891 -3322	여	1983. 12.3.	경기도 용인시 죽전동 33-11	OO 법무법인	7,000	1	0	기업 은행
…	…	…	…	…	…		…	…			…

나. 재식별 의도·재식별 가능성 판단에 따른 가명처리

○ 자체 보유 개인신용정보를 가명처리하여 분석하고자 하는 상황으로, 가명정보를 생성한 회사의 내부에서 가명정보가 활용되고 있어 가명정보를 불법적으로 재식별할 의도가 상대적으로 낮다고 볼 수 있음

ㅇ 가명정보에 대해 개인신용정보에 준하는 보안대책을 적용하여 내부 통제하에 활용하고 있으므로 재식별 가능성도 낮게 판단됨

※ 가명처리의 수준을 결정하기 위한 가명정보의 재식별 위험도 측정시 해당 이용기관의 재식별 의도와 능력, 가명정보 보호능력, 업무수행 신뢰도 등 다양한 측면에서 평가하여 이를 종합적으로 고려

< (예시) 속성자 분류 및 가명처리 >

구분	속성	위험	가명처리 예시*
ID	식별자	신용카드사에서 개인을 식별하기 위한 ID이므로 개인이 특정될 가능성이 있음. 단, 시계열 분석시 동일인 여부를 확인할 수 있는 값으로서 가명처리 후 활용	랜덤값 생성후 대체 (매핑테이블에 대해서는 분리보관 조치)
성명		식별자로서 개인을 특정할 수 있음	삭제
카드번호		식별자이며 다른 사업자도 보유하고 있을 수 있으므로 개인이 식별될 위험이 존재	삭제
전화번호		식별자이며 다른 사업자도 보유하고 있을 수 있으므로 개인이 식별될 위험이 존재	삭제
성별	개인식별 가능정보	생년월일, 주소 등의 정보와 조합될 경우 개인이 식별될 수 있음	별도 조치 없이 활용
생년월일		주소, 성별 등의 정보와 조합하여 개인이 식별될 수 있음	별도 조치 없이 활용
주소		상세주소의 경우 그 자체로 식별자로 판단될 수 있으며, 다른 정보와 조합하여 개인이 식별될 위험이 존재	별도 조치 없이 활용
직업 (근무처)		공인 등 일부 직업의 경우 다른 정보와 조합하여 개인이 식별될 위험이 존재	특이치 존재시 상·하단 코딩 적용
연봉		수입이 너무 많거나 적은 경우 다른 정보와 조합하여 개인이 식별될 위험이 존재	특이치 존재시 상·하단 코딩 적용
내부신용 등급		이미 등급화된 정보이며 재식별될 경우 개인에 민감한 정보	별도 조치 없이 활용
연체잔고		외부에 유출될 경우 개인에 민감한 정보	특이치 존재시 상·하단 코딩 적용
결제기관		결제기관 정보로는 개인이 식별될 위험이 크지 않음	별도 조치 없이 활용

* 가명처리 목적, 처리·이용환경, 이용 주체 등에 따른 위험도(Risk) 분석을 통해 결정

[가명처리 예시 ②] 내부 결합

▷ 처리 목적 : A카드회사는 2021년 상반기 출시 예정인 새로운 중금리 대출 상품을 준비하면서 대출심사전략을 마련하고자 2019년 카드이용 고객 중 중금리 상품 이용 예상 고객을 추출하여 관련 분석·연구를 추진(**동일 회사 내 다른 부서간 데이터 결합하여 분석 예정)**

▷ 가명정보 활용구간 : 내부

▷ 보호조치·접근통제 수준 : 가명정보 및 추가정보에 대해 내부관리계획에 따라 개인신용정보에 준하는 기술적·관리적·물리적 보안 대책을 적용하고 내부통제 하에 활용

가. 보유 정보 샘플

ID	성명	카드 번호	전화 번호	성별	생년 월일	주소	근무처	연봉 (만원)	내부 신용 등급	연체 잔고 (만원)	결제 기관
19342	홍길동	3779 4593 3043 3921 3943	010 -3355 -0934	남	1972. 9.9.	서울시 강남구 역삼동 332-1	OO 자동차	4,500	5	35	국민 은행
19354	김철수	4832 2332 2344 4399	02 -531 -9834	남	1980. 4.16.	서울시 마포구 송내동 334-1	OO 은행	6,000	2	0	새마을 금고
19445	전지연	4523 3234 9843 0394	010 -9290 -3344	여	1979. 5.23.	경기도 광주시 송정동 786-1	OO 공사	5,500	9	125	우리 은행
20221	박식별	4932 3453 5943 4321	010 -2891 -3322	여	1983. 12.3.	경기도 용인시 죽전동 33-11	OO 법무법인	7,000	1	0	기업 은행
…	…	…	…	…	…		…	…			…

나. 재식별 의도·재식별 가능성 판단에 따른 가명처리

○ 자체 보유하고 있는 개인신용정보를 가명처리하여 분석하고자 하는 상황으로, 가명정보를 생성한 회사의 내부에서 가명정보가 활용되고 있어 가명정보를 불법적으로 재식별 할 의도가 상대적으로 낮음

○ 가명정보에 대해 개인신용정보에 준하는 보안대책을 적용하고 내부 통제하에 활용하고 있어 재식별 가능성도 낮게 판단되나, **서로 다른 부서(업무)의 데이터를 결합하여 활용하는 과정에서 속성자 간의 조합을 통해 재식별되지 않도록 주의 필요**

※ 가명처리의 수준을 결정하기 위한 가명정보의 재식별 위험도 측정시 해당 이용기관의 재식별 의도와 능력, 가명정보 보호능력, 업무수행 신뢰도 등 다양한 측면에서 평가하여 이를 종합적으로 고려

〈(예시) 속성자 분류 및 가명처리〉

구분	속성	위험	가명처리 예시*
ID	식별자	신용카드사에서 개인을 식별하기 위한 ID이므로 개인이 특정될 가능성이 있음. 단, 시계열 분석시 동일인 여부를 확인할 수 있는 값으로서 가명처리 후 활용	랜덤값 생성후 대체 (매핑테이블에 대해서는 분리보관 조치)
성명		식별자로서 개인을 특정할 수 있음	삭제
카드번호		식별자이며 다른 사업자도 보유하고 있을 수 있으므로 개인이 식별될 위험이 존재	삭제
전화번호		식별자이며 다른 사업자도 보유하고 있을 수 있으므로 개인이 식별될 위험이 존재	삭제
성별	개인식별 가능정보	생년월일, 주소 등의 정보와 조합될 경우 개인이 식별될 수 있음	코드값으로 대체
생년월일		주소, 성별 등의 정보와 조합하여 개인이 식별될 수 있음	별도 조치 없이 활용 (다른 개인식별가능정보를 조치하여 위험 해소)
주소		상세주소의 경우 그 자체로 식별자로 판단될 수 있으며, 다른 정보와 조합하여 개인이 식별될 위험이 존재	분석 목적상 세부주소는 불필요하므로, 동단위 하위 주소 삭제
직업 (근무처)		공인 등 일부 직업의 경우 다른 정보와 조합하여 개인이 식별될 위험이 존재	별도 조치 없이 활용 (다른 개인식별가능정보를 조치하여 위험 해소)
연봉		수입이 너무 많거나 적은 경우 다른 정보와 조합하여 개인이 식별될 위험이 존재	특이치 존재시 상·하단 코딩 적용
내부신용 등급		이미 등급화된 정보이며 외부에 유출될 경우 개인에 민감한 정보	별도 조치 없이 활용
연체잔고		외부에 유출될 경우 개인에 민감한 정보	특이치 존재시 상·하단 코딩 적용
결제기관		결제기관 정보로는 개인이 식별될 위험이 크지 않음	별도 조치 없이 활용

* 가명처리 목적, 처리·이용환경, 이용 주체 등에 따른 위험도(Risk) 분석을 통해 결정

[가명처리 예시 ③] 외부기관간 결합

▶ 처리 목적 : A카드회사는 2021년 상반기 출시 예정인 새로운 중금리 대출 상품을 준비하면서 대출심사전략을 마련하고자 2019년 A사 카드 이용고객 데이터와 B신용평가사의 신용정보조회고객 데이터를 결합하여 관련 분석·연구를 추진

▶ 가명정보 활용구간 : 내부

▶ 보호조치·접근통제 수준 : 가명정보 및 추가정보에 대해 내부관리계획에 따라 개인신용정보에 준하는 기술적·관리적·물리적 보안 대책을 적용하고 내부통제 하에 활용

가. 보유 정보 샘플

ID	성명	카드 번호	전화 번호	성별	생년 월일	주소	근무처	연봉 (만원)	내부 신용 등급	연체 잔고 (만원)	결제 기관
19342	홍길동	3779 4593 3043 3921 3943	010 -3355 -0934	남	1972. 9.9.	서울시 강남구 역삼동 332-1	OO 자동차	4,500	5	35	국민 은행
19354	김철수	4832 2332 2344 4399	02 -531 -9834	남	1980. 4.16.	서울시 마포구 송내동 334-1	OO 은행	6,000	2	0	새마을 금고
19445	전지연	4523 3234 9843 0394	010 -9290 -3344	여	1979. 5.23.	경기도 광주시 송정동 786-1	OO 공사	5,500	9	125	우리 은행
20221	박식별	4932 3453 5943 4321	010 -2891 -3322	여	1983. 12.3.	경기도 용인시 죽전동 33-11	OO 법무법인	7,000	1	0	기업 은행
…	…	…	…	…	…		…	…			…

나. 재식별 의도·재식별 가능성 판단에 따른 가명처리

○ 자체 보유 개인신용정보를 가명처리하여 타 회사 데이터와 결합하여 활용하고자 하는 상황으로, 가명정보를 생성한 회사의 내부에서 가결합된 가명정보가 활용되고 있어 가명정보를 불법적으로 재식별할 의도가 상대적으로 낮다고 볼 수 있음

○ 가명정보에 대해 개인신용정보에 준하는 보안대책을 적용하여 내부 통제하에 활용하고 있으므로 재식별 가능성도 낮게 판단됨

○ 단, 타 회사의 데이터와 결합된 결합정보의 경우 의도하지 않게 재식별될 수도 있으므로 [가명처리 예시①]보다는 가명처리 수준을 강화할 필요 있음

※ 가명처리의 수준을 결정하기 위한 가명정보의 재식별 위험도 측정시 해당 이용기관의 재식별 의도와 능력, 가명정보 보호능력, 업무수행 신뢰도 등 다양한 측면에서 평가하여 이를 종합적으로 고려

< (예시) 속성자 분류 및 가명처리 >

구분	속성	위험	가명처리 예시*
ID	식별자	신용카드사에서 개인을 식별하기 위한 ID이므로 개인이 특정될 가능성이 있음.	삭제
성명		식별자로서 개인을 특정할 수 있음	삭제
카드번호		식별자이며 다른 사업자도 보유하고 있을 수 있으므로 개인이 식별될 위험이 존재	삭제
전화번호		식별자이며 다른 사업자도 보유하고 있을 수 있으므로 개인이 식별될 위험이 존재	단방향 암호화 처리 (결합키로 활용)
성별	개인식별 가능정보	생년월일, 주소 등의 정보와 조합될 경우 개인이 식별될 수 있음	별도 조치 없이 활용 (다른 개인식별가능정보를 조치하여 위험 해소)
생년월일		주소, 성별 등의 정보와 조합하여 개인이 식별될 수 있음	연령대로 변환
주소		상세주소의 경우 그 자체로 식별자로 판단될 수 있으며, 다른 정보와 조합하여 개인이 식별될 위험이 존재	분석 목적상 세부주소는 불필요하므로, 동단위 하위 주소 삭제
직업 (근무처)		공인 등 일부 직업의 경우 다른 정보와 조합하여 개인이 식별될 위험이 존재	제조업, 금융업 등 직종으로 일반화
연봉		수입이 너무 많거나 적은 경우 다른 정보와 조합하여 개인이 식별될 위험이 존재	특이치 존재시 상·하단 코딩 적용
내부신용 등급		이미 등급화된 정보이며 외부에 유출될 경우 개인에 민감한 정보	별도 조치 없이 활용
연체잔고		외부에 유출될 경우 개인에 민감한 정보	특이치 존재시 상·하단 코딩 적용
결제기관		결제기관 정보로는 개인이 식별될 위험이 크지 않음	별도 조치 없이 활용

* 가명처리 목적, 처리·이용환경, 이용 주체 등에 따른 위험도(Risk) 분석을 통해 결정

[가명처리 예시 ④] 외부 제공

▶ 처리 목적 : C대학의 연구실은 A카드회사가 보유한 개인신용정보 10년치를 분석하여 2008년 금융위기 이후 신용카드 이용자의 연령대, 거주지역, 근무지역, 직종, 연봉 등과 소비의 상관관계를 연구하고자 함

▶ 가명정보 활용구간 : 외부(대학 연구실)

▶ 보호조치·접근통제 수준 : C대학 연구실의 기술적·관리적·물리적 보안수준은 신용카드회사 보다 낮은 것으로 확인되었음

가. 보유 정보 샘플

ID	성명	카드 번호	전화 번호	성별	생년 월일	주소	근무처	연봉 (만원)	내부 신용 등급	월평균 결제액 (만원)	결제 기관
19342	홍길동	3779 4593 3043 3921 3943	010 -3355 -0934	남	1972. 9.9.	서울시 강남구 역삼동 332-1	OO 자동차	4,500	5	35	국민 은행
19354	김철수	4832 2332 2344 4399	02 -531 -9834	남	1980. 4.16.	서울시 마포구 송내동 334-1	OO 은행	6,000	2	240	새마을 금고
19445	전지연	4523 3234 9843 0394	010 -9290 -3344	여	1979. 5.23.	경기도 광주시 송정동 786-1	OO 공사	5,500	9	125	우리 은행
20221	박식별	4932 3453 5943 4321	010 -2891 -3322	여	1983. 12.3.	경기도 용인시 죽전동 33-11	OO 법무법인	7,000	1	85	기업 은행
…	…	…	…	…	…		…	…			…

나. 재식별 의도·재식별 가능성 판단에 따른 가명처리

○ 자체 보유한 개인신용정보를 가명처리하여 연구실에 제공하고자 하는 상황으로, 가명정보를 생성한 카드사 외부로 가명정보가 반출되어 활용되고 가명정보를 활용하는 연구실의 보안수준이 카드사보다 낮은 것으로 확인되므로 개인신용정보가 재식별 되지 않도록 각별한 주의 필요

※ 가명처리의 수준을 결정하기 위한 가명정보의 재식별 위험도 측정시 해당 이용기관의 재식별 의도와 능력, 가명정보 보호능력, 업무수행 신뢰도 등 다양한 측면에서 평가하여 이를 종합적으로 고려

< (예시) 속성자 분류 및 가명처리 >

구분	속성	위험	가명처리 예시*
ID	식별자	신용카드사에서 개인을 식별하기 위한 ID이므로 개인이 특정될 가능성이 있음.	삭제
성명		식별자로서 개인을 특정할 수 있음	삭제
카드번호		식별자이며 다른 사업자도 보유하고 있을 수 있으므로 개인이 식별될 위험이 존재	삭제
전화번호		식별자이며 다른 사업자도 보유하고 있을 수 있으므로 개인이 식별될 위험이 존재	삭제
성별	개인식별 가능정보	생년월일, 주소 등의 정보와 조합될 경우 개인이 식별될 수 있음	코드값으로 대체
생년월일		주소, 성별 등의 정보와 조합하여 개인이 식별될 수 있음	연령대로 변환
주소		상세주소의 경우 그 자체로 식별자로 판단될 수 있으며, 다른 정보와 조합하여 개인이 식별될 위험이 존재	분석 목적상 세부주소는 불필요하므로, 동단위 하위 주소 삭제
직업 (근무처)		공인 등 일부 직업의 경우 다른 정보와 조합하여 개인이 식별될 위험이 존재	제조업, 금융업 등 직종으로 일반화
연봉		수입이 너무 많거나 적은 경우 다른 정보와 조합하여 개인이 식별될 위험이 존재	특이치 존재시 상·하단 코딩 적용
내부신용 등급		이미 등급화된 정보이며 외부에 유출될 경우 개인에 민감한 정보	분석 목적상 필요하므로, 별도 조치 없이 활용
월평균 결제금액		외부에 유출될 경우 개인에 민감한 정보	범주화하여 활용 특이치는 레코드 삭제
결제기관		결제기관 정보로는 개인이 식별될 위험이 크지 않음	별도 조치 없이 활용

* 가명처리 목적, 처리·이용환경, 이용 주체 등에 따른 위험도(Risk) 분석을 통해 결정

4. 가명처리에 관한 행위규칙

※「신용정보법」 제40조의2(가명처리·익명처리에 관한 행위규칙),「신용정보업감독규정」[별표 8] 가명정보에 관한 보호조치 기준(본 안내서 'Ⅱ. 5. 가명정보 및 추가정보에 관한 보호조치 기준' 참조) 등 참고

가. 추가정보의 분리 보관 또는 삭제

신용정보회사등은 가명처리에 사용한 추가정보를 분리하여 보관하거나 삭제하여야 한다. 이 때 금융위원회가 정하여 고시하는 기술적·관리적·물리적 보호조치를 통해 추가정보에 접근하는 것을 통제하는 방법을 준수하여야 한다.

나. 기술적 · 관리적 · 물리적 보안대책 수립 · 시행

신용정보회사등은 가명처리한 개인신용정보에 대하여 제3자의 불법적인 접근, 입력된 정보의 변경·훼손 및 파괴, 그 밖의 위험으로부터 가명정보를 보호하기 위하여 내부관리계획을 수립하고 접속기록을 보관하는 등 기술적 · 관리적 · 물리적 보안대책을 수립 · 시행하여야 하며, 다음 사항을 포함하여야 한다.

1) 가명처리한 개인신용정보에 제3자가 불법적으로 접근하는 것을 차단하기 위한 침입차단시스템 등 접근통제장치의 설치 · 운영에 관한 사항
2) 가명처리한 개인신용정보의 변경 · 훼손 및 파괴를 방지하기 위한 사항
3) 가명처리한 개인신용정보 취급 · 조회 권한을 직급별 · 업무별로 차등 부여하는 데에 관한 사항 및 가명처리한 개인신용정보 접근기록의 주기적인 점검에 관한 사항
4) 가명처리전 개인신용정보와 가명처리한 개인신용정보의 분리에 관한 사항
5) 가명정보를 통계작성, 연구, 공익적 기록보존 등의 목적으로 이용 · 제공할 경우 해당 목적 외 활용 방지에 관한 사항
6) 그 밖에 가명처리한 개인신용정보의 안전성 확보를 위하여 금융위원회가 정하여 고시하는 사항

다. 가명처리의 제한

신용정보회사등은 영리 또는 부정한 목적으로 특정 개인을 알아볼 수 있게 가명정보를 처리하여서는 아니 된다.

라. 재식별시 조치

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.

신용정보회사등은 가명정보 재식별 이력이 있을 경우 신용정보관리·보호인에게 보고하고 기록·관리할 수 있도록 하여야 한다.

마. 가명처리 기록의 보존

신용정보회사등은 개인신용정보를 가명처리 한 경우 다음의 항목을 포함하여 그 조치 기록을 3년간 보존하여야 한다.

1) 가명처리한 날짜
2) 가명처리한 정보의 항목
3) 가명처리한 사유와 근거

< (예시) 가명처리 기록 >

날짜	가명처리 근거	정보항목	가명처리 사유	가명처리 방법
2020. 9. 1.	① 가명처리 목적 연령대별 신용등급에 따른 연체율 연구 ② 관련 문서 첨부 (연구 계획(안) 등) ③ 근거 규정 「신용정보법」 제32조 제6항 제9호의2 등	고객ID	식별자	삭제
		이름	식별자	이름, 휴대폰번호를 조합하여 해시함수(SHA-256, 솔트값 적용)로 ID생성 후 삭제
		휴대폰번호	식별자	
		나이	개인식별가능정보 (재식별 가능성이 있고 연구목적상 구체적인 나이는 불필요)	연령대로 범주화
		대출금액	개인식별가능정보 (연구목적상 구체적인 수치는 불필요)	만원 단위로 반올림
		연체기록	개인식별가능정보 (구체적인 수치는 재식별 우려가 있음)	연체여부(Y/N)만 표기
		신용등급	개인식별가능정보 (신용등급은 이미 범주화된 등급이어서 재식별 우려가 거의 없음)	별도 조치 없음
		…	…	…
2020. 11. 3.	① 가명처리 목적 보험가입자 특성에 대한 공동연구(X사의 데이터와 결합 후 분석) ② 관련 문서 첨부 (양사간 계약서, 공동연구 계획(안) 등) ③ 근거 규정 「신용정보법」 제17조의2, 제32조제6항 제9호의2, 동법 시행령 제14조의2 등	고객ID	식별자	삭제
		이름	식별자	이름, 휴대폰번호를 조합하여 해시함수(SHA-256, 솔트값 적용)로 ID생성 후 삭제
		휴대폰번호	식별자	
		거래지점	개인식별가능정보 (구체적인 지점 정보는 재식별 우려가 있음)	구 단위로 범주화
		보험가입건수	개인식별가능정보 (구체적인 가입건수 수치는 재식별 우려가 낮음)	특이치만 삭제하고 별도 조치 없이 활용
		약관대출금액	개인식별가능정보 (연구목적 상 구체적인 수치는 불필요)	십만원 단위로 반올림
		…	…	…

바. 가명처리 관련 사항의 공개

가명정보는 통계작성, 연구, 공익적 기록보존 등 제한된 목적 내에서만 활용 가능한 특수한 형태의 개인신용정보이다. 따라서 가명정보를 처리하는 기관은 가명처리 관련 사항을 「신용정보법」, 「개인정보 보호법」 등과 같은 법률에 따라 공개하여야 한다.

신용정보회사, 신용정보집중기관 및 동법 시행령 제27조에서 정하는 신용정보제공·이용자는 가명정보 활용과 관련된 사항을 신용정보활용체제에 포함하여 공시하여야 한다(「신용정보법」 제31조).

< (예시) 가명정보 처리 관련 신용정보활용체제에 포함될 사항 >

1. 개인신용정보의 가명처리 관련 사항을 포함된 개인신용정보 보호 및 관리에 관한 기본계획
2. 처리하는 가명정보의 종류 및 이용 목적
3. 가명정보를 제3자에게 제공하는 경우 제공하는 가명정보의 종류, 제공 대상, 제공받는 자의 이용 목적
4. 가명정보의 보유 또는 이용 기간, 가명정보 파기의 절차 및 방법
5. 가명정보 처리의 위탁이 있는 경우 그 업무의 내용 및 수탁자
6. 처리하는 가명정보의 항목

◎ **「신용정보법」 제31조(신용정보활용체제의 공시)** ① 개인신용평가회사, 개인사업자신용평가회사, 기업신용조회회사, 신용정보집중기관 및 대통령령으로 정하는 신용정보제공·이용자는 다음 각 호의 사항을 대통령령으로 정하는 바에 따라 공시하여야 한다.

1. 개인신용정보 보호 및 관리에 관한 기본계획(총자산, 종업원 수 등을 고려하여 대통령령으로 정하는 자로 한정한다)
2. 관리하는 신용정보의 종류 및 이용 목적
3. 신용정보를 제공받는 자
4. 신용정보주체의 권리의 종류 및 행사 방법
5. 신용평가에 반영되는 신용정보의 종류, 반영비중 및 반영기간(개인신용평가회사, 개인사업자신용평가회사 및 기업신용등급제공업무·기술신용평가업무를 하는 기업신용조회회사로 한정한다)

6.「개인정보 보호법」 제30조제1항제6호 및 제7호의 사항
7. 그 밖에 신용정보의 처리에 관한 사항으로서 대통령령으로 정하는 사항

◎「신용정보법」 시행령

제27조(신용정보활용체제의 공시) ① 법 제31조에서 "대통령령으로 정하는 신용정보제공·이용자"란 제5조제1항제1호부터 제21호까지 및 제21조제2항제1호부터 제21호까지의 규정의 어느 하나에 해당하는 기관을 말한다.
② 신용정보회사, 신용정보집중기관 및 제1항에 해당하는 자는 법 제31조에 따라 다음 각 호의 사항을 공시하여야 한다.
1. 관리하는 신용정보의 종류 및 이용 목적
2. 신용정보를 제3자에게 제공하는 경우 제공하는 신용정보의 종류, 제공 대상, 제공받는 자의 이용 목적(제1항에 해당하는 자로 한정한다)
3. 신용정보의 보유 기간 및 이용 기간이 있는 경우 해당 기간, 신용정보 파기의 절차 및 방법(제1항에 해당하는 자로 한정한다)
4. 법 제17조에 따라 신용정보의 처리를 위탁하는 경우 그 업무의 내용 및 수탁자
5. 신용정보주체의 권리와 그 행사방법
6. 법 제20조제3항에 따른 신용정보관리·보호인 또는 신용정보 관리·보호 관련 고충을 처리하는 사람의 성명, 부서 및 연락처
7. 신용등급 산정에 반영되는 신용정보의 종류, 반영비중 및 반영기간(신용조회회사만 해당한다)

제5조(신용정보업별 허가 대상) ① 법 제5조제1항제1호에서 "대통령령으로 정하는 금융기관"이란 다음 각 호의 어느 하나에 해당하는 기관을 말한다. 다만, 제9호부터 제14호까지의 경우에는 그 연합회 또는 중앙회만 말한다.
1.「은행법」에 따라 인가를 받아 설립된 은행(같은 법 제59조에 따라 은행으로 보는 자를 포함한다)
2.「금융지주회사법」에 따른 금융지주회사
3.「한국산업은행법」에 따른 한국산업은행
4.「한국수출입은행법」에 따른 한국수출입은행
5.「농업협동조합법」 제161조의11에 따른 농협은행
5의2.「수산업협동조합법」에 따른 수협은행
6.「중소기업은행법」에 따른 중소기업은행
7.「한국주택금융공사법」에 따른 한국주택금융공사
8.「자본시장과 금융투자업에 관한 법률」에 따른 금융투자업자·증권금융회사·종합금융회사·자금중개회사 및 명의개서대행회사
9.「상호저축은행법」에 따른 상호저축은행과 그 중앙회

10. 「농업협동조합법」에 따른 농업협동조합과 그 중앙회
11. 「수산업협동조합법」에 따른 수산업협동조합과 그 중앙회
12. 「산림조합법」에 따른 산림조합과 그 중앙회
13. 「신용협동조합법」에 따른 신용협동조합과 그 중앙회
14. 「새마을금고법」에 따른 새마을금고와 그 연합회
15. 「보험업법」에 따른 보험회사
16. 「여신전문금융업법」에 따른 여신전문금융회사(「여신전문금융업법」 제3조제3항제1호에 따라 허가를 받거나 등록을 한 자를 포함한다)
17. 「기술보증기금법」에 따른 기술보증기금
18. 「신용보증기금법」에 따른 신용보증기금
19. 「지역신용보증재단법」에 따른 신용보증재단과 그 중앙회
20. 「무역보험법」에 따른 한국무역보험공사
21. 「예금자보호법」에 따른 예금보험공사 및 정리금융회사

제21조(신용정보의 집중관리·활용) ② 법 제25조제2항제1호에서 "대통령령으로 정하는 금융기관"이란 제5조제1항제1호부터 제20호까지의 금융기관 및 다음 각 호에 해당하는 기관을 말한다.

1. 「건설산업기본법」에 따른 공제조합
2. 「국채법」에 따른 국채등록기관
3. 「한국농수산식품유통공사법」에 따른 한국농수산식품유통공사
4. 「서민의 금융생활 지원에 관한 법률」 제56조에 따른 신용회복위원회
5. 「산업재해보상보험법」에 따른 근로복지공단
6. 「소프트웨어산업 진흥법」에 따른 소프트웨어공제조합
7. 「엔지니어링산업 진흥법」에 따른 엔지니어링공제조합
8. 「예금자보호법」에 따른 정리금융회사
9. 「우체국예금·보험에 관한 법률」에 따른 체신관서
10. 「전기공사공제조합법」에 따른 전기공사공제조합
11. 「주택도시기금법」에 따른 주택도시보증공사
12. 「중소기업진흥에 관한 법률」에 따른 중소벤처기업진흥공단
13. 「중소기업창업 지원법」에 따른 중소기업창업투자회사 및 중소기업창업투자조합
14. 「중소기업협동조합법」에 따른 중소기업중앙회
15. 「한국장학재단 설립 등에 관한 법률」에 따른 한국장학재단
16. 「금융회사부실자산 등의 효율적 처리 및 한국자산관리공사의 설립에 관한 법률」에 따른 한국자산관리공사
17. 「상법」에 따라 설립된 주식회사 국민행복기금
18. 「서민의 금융생활 지원에 관한 법률」 제3조에 따른 서민금융진흥원

19.「대부업 등의 등록 및 금융이용자 보호에 관한 법률」 제3조제2항에 따라 금융위원회에 등록한 대부업자등
20.「산업발전법」 제40조제1항제1호에 따른 자본재공제조합
21.「소상공인 보호 및 지원에 관한 법률」 제17조제1항에 따른 소상공인시장진흥공단

또한 개인정보처리자는 가명정보의 처리와 관련된 사항을 포함하여 개인정보처리방침을 작성 후 공개하여야 한다(「개인정보 보호법」 제30조).

◎ **「개인정보 보호법」 제30조(개인정보 처리방침의 수립 및 공개)** ① 개인정보처리자는 다음 각 호의 사항이 포함된 개인정보의 처리 방침(이하 "개인정보 처리방침"이라 한다)을 정하여야 한다. 이 경우 공공기관은 제32조에 따라 등록대상이 되는 개인정보파일에 대하여 개인정보 처리방침을 정한다.
1. 개인정보의 처리 목적
2. 개인정보의 처리 및 보유 기간
3. 개인정보의 제3자 제공에 관한 사항(해당되는 경우에만 정한다)
3의2. 개인정보의 파기절차 및 파기방법(제21조제1항 단서에 따라 개인정보를 보존하여야 하는 경우에는 그 보존근거와 보존하는 개인정보 항목을 포함한다)
4. 개인정보처리의 위탁에 관한 사항(해당되는 경우에만 정한다)
5. 정보주체와 법정대리인의 권리·의무 및 그 행사방법에 관한 사항
6. 제31조에 따른 개인정보 보호책임자의 성명 또는 개인정보 보호업무 및 관련 고충사항을 처리하는 부서의 명칭과 전화번호 등 연락처
7. 인터넷 접속정보파일 등 개인정보를 자동으로 수집하는 장치의 설치·운영 및 그 거부에 관한 사항(해당하는 경우에만 정한다)
8. 그 밖에 개인정보의 처리에 관하여 대통령령으로 정한 사항

사. 가명정보에 대한 적용 예외

가명처리된 개인신용정보도 「신용정보법」상의 개인신용정보에 해당하나 동법은 개인신용정보에 적용되는 규정 중 보유기간, 개인신용정보 제공·활용에 대한 동의 등 일부에 대해서는 적용 예외를 두는 등 달리 정하고 있다.

◎「신용정보법」

제20조의2(개인신용정보의 보유기간 등) ②「개인정보 보호법」 제21조제1항에도 불구하고 신용정보제공·이용자는 금융거래 등 상거래관계가 종료된 날부터 최장 5년 이내(해당 기간 이전에 정보 수집·제공 등의 목적이 달성된 경우에는 그 목적이 달성된 날부터 3개월 이내)에 해당 신용정보주체의 개인신용정보를 관리 대상에서 삭제하여야 한다. 다만, 다음 각 호의 경우에는 그러하지 아니하다.
2의2. 가명정보를 이용하는 경우로서 그 이용 목적, 가명처리의 기술적 특성, 정보의 속성 등을 고려하여 대통령령으로 정하는 기간 동안 보존하는 경우

제40조의3(가명정보에 대한 적용 제외) 가명정보에 관하여는 제32조제7항, 제33조의2, 제35조, 제35조의2, 제35조의3, 제36조, 제36조의2, 제37조, 제38조, 제38조의2, 제38조의3, 제39조 및 제39조의2부터 제39조의4까지의 규정을 적용하지 아니한다.

제32조(개인신용정보의 제공·활용에 대한 동의)
제33조의2(개인신용정보의 전송요구)
제35조(신용정보 이용 및 제공사실의 조회)
제35조의2(개인신용평점 하락 가능성 등에 대한 설명의무)
제35조의3(신용정보제공·이용자의 사전통지)
제36조(상거래 거절 근거 신용정보의 고지 등)
제36조의2(자동화평가 결과에 대한 설명 및 이의제기 등)
제37조(개인신용정보 제공 동의 철회권 등)
제38조(신용정보의 열람 및 정정청구 등)
제38조의2(신용조회사실의 통지 요청)
제38조의3(개인신용정보의 삭제 요구)
제39조(무료 열람권)
제39조의2(채권자변동정보의 열람 등)
제39조의3(신용정보주체의 권리행사 방법 및 절차)
제39조의4(개인신용정보 누설통지 등)

5. 가명정보 및 추가정보에 관한 보호조치 기준

※「신용정보업감독규정」[별표 8] 가명정보에 관한 보호조치 기준(제43조의7 관련)

가. 기술적·물리적 보호조치

1) 추가정보에 대한 보호조치

가) 신용정보회사등은 추가정보를 삭제하지 아니하고 보존하여야 하는 경우 추가정보를 가명정보와 분리된 저장소*에 암호화하여 저장하여야 한다.

※ 반드시 추가정보를 보존하여야 하는 경우를 제외하고는 추가정보를 삭제할 것을 권고

* 논리적·물리적 분리방법 모두 가능하나, 테이블을 분리하는 방법은 허용되지 않음

나) 신용정보회사등은 원칙적으로 가명정보를 취급하는 직원이 추가정보에 접근할 수 있는 권한을 부여하지 않아야 하며, 추가정보 접근이 불가피한 경우 관리책임자의 사전 승인을 받아 일시적으로 부여하고 관련 기록을 보관하는 등 적절한 통제시스템을 갖추어야 한다.

다) 신용정보회사등은 위 '나)'에 따른 기록 보관시 접근자의 신원, 관리책임자의 신원, 접근일시, 대상정보, 조회가 불가피한 사유, 용도 등의 기록을 3년간 보관하여야 한다.

라) 신용정보회사등은 추가정보가 가명정보를 재식별하는 데 사용되는 등 부정한 목적으로 사용되지 않도록 월 1회 이상 주기적으로 점검하여야 한다.

◎ 추가정보에 관한 보호조치를 강화하기 위해,「신용정보법」및 동법 시행령,「신용정보업감독규정」상의 의무사항은 아니나 **원본정보와 추가정보의 담당자도 분리하는 방안도 고려할 수 있음**

2) 가명정보에 대한 보호조치

가) 신용정보회사등은 가명처리전 개인신용정보와 가명처리한 개인신용정보를 분리하여 저장하여야 한다.

나) 신용정보회사등은 가명정보를 취급하는 담당자를 별도로 지정·관리하고 가명처리전 개인신용정보를 취급하는 담당자와 접근권한을 구분하여 운영하여야 한다.

다) 신용정보회사등은 원칙적으로 가명정보를 취급하는 직원이 가명처리전 개인신용정보에 접근할 수 있는 권한을 부여하지 않아야 하며, 원본정보 접근이 불가피한 경우 관리책임자의 사전 승인을 득하여 일시적으로 부여하고, 관련 기록을 보관하는 등 적절한 통제 시스템을 갖추어야 한다.

라) 신용정보회사등은 위 ‘다)’에 따른 기록 보관시 접근자의 신원, 관리책임자의 신원, 접근일시, 대상정보, 접근이 불가피한 사유, 용도 등의 기록을 3년 이상 보관하여야 한다.

마) 신용정보회사등은 가명정보 처리 시 가명정보의 구체적인 처리 목적, 처리 방법, 처리 일시를 기록하여 가명정보가 파기된 이후 3년 이상 보관하고, 처리 기록에 대해 월 1회 이상 주기적으로 확인·감독하여야 한다.

바) 신용정보회사등은 가명정보 오·남용에 대한 자체 제재기준을 마련하여야 한다.

나. 관리적 보호조치

1) 신용정보회사등은 가명처리한 개인신용정보에 대하여 제3자의 불법적인 접근, 입력된 정보의 변경·훼손 및 파괴, 그 밖의 위험으로부터 가명정보를 보호하기 위해 다음 각 호의 사항을 포함하는 별도 내부관리계획을 수립·시행하여야 한다.

가) 가명정보 및 추가정보에 대한 접근 권한 부여·변경·말소에 관한 사항
나) 가명정보 및 추가정보가 저장 또는 처리되는 시스템・단말의 보호조치에 관한 사항
다) 가명정보 및 추가정보에 대한 접근기록 보관 및 점검에 관한 사항
라) 가명정보 및 추가정보의 보유 기간 및 파기 기준・방법에 관한 사항
마) 가명정보의 목적 외 활용 방지 및 재식별 방지 대책에 관한 사항
바) 가명정보 제3자 제공 시 사후관리에 관한 사항

2) 신용정보회사등은 가명정보 및 추가정보에 접근하는 취급자들에 대해 다음의 사항을 포함하는 가명정보보호교육을 연 1회 이상 수행하여야 한다.

가) 가명정보의 목적 외 활용 금지에 관한 사항
나) 가명정보의 재식별 금지에 관한 사항
다) 가명정보 재식별 시 즉시 회수 및 삭제에 관한 사항

3) 신용정보회사등은 다음의 사항을 고려하여 가명정보의 보존기간을 주기적으로 검토하고, 그 적정성 여부를 판단하여 필요시 조정하여야 한다.

가) 추가정보 및 가명정보에 대한 기술적・관리적・물리적 보호조치 수준
나) 가명정보의 재식별시 정보주체에 미치는 영향
다) 가명정보의 재식별 가능성
라) 가명정보의 이용목적 및 그 목적 달성에 필요한 최소기간

4) 신용정보회사등은 가명정보를 통계작성, 연구, 공익적 기록보존 등을 위하여 제공하는 경우 개인신용정보의 제공·활용에 대한 동의에 관련 의무(「신용정보법」 제32조제1항 내지 제5항)가 적용되지 않는다(동법 제32조제6항 제9호의2). 제3자에게 제공하는 경우 다음의 사항을 준수하여야 한다.

가) 가명정보를 불특정 다수에게 공개하지 아니할 것
나) 가명정보 제공 시 가명정보를 제공 받는 자, 가명정보 활용목적, 가명정보 이용·보존기간 등을 구체적으로 명시하여 제공할 것
다) 가명정보의 재식별 금지, 가명정보의 목적 외 사용 금지 등 관련 법령 준수에 관한 사항을 주지시킬 것
라) 추가정보를 제공하거나 공개하지 않을 것
마) 가명정보의 재식별 가능성을 발견한 경우에는 즉시 그 정보를 처리하고 있는 자에게 통지하고 처리중단 요구 및 해당정보를 회수·파기하는 조치를 취할 것

◎ **「신용정보법」 제32조(개인신용정보의 제공·활용에 대한 동의)** ① 신용정보제공·이용자가 개인신용정보를 타인에게 제공하려는 경우에는 대통령령으로 정하는 바에 따라 해당 신용정보주체로부터 다음 각 호의 어느 하나에 해당하는 방식으로 개인신용정보를 제공할 때마다 미리 개별적으로 동의를 받아야 한다. 다만, 기존에 동의한 목적 또는 이용 범위에서 개인신용정보의 정확성·최신성을 유지하기 위한 경우에는 그러하지 아니하다.

1. 서면
2. 「전자서명법」 제2조제3호에 따른 공인전자서명이 있는 전자문서(「전자문서 및 전자거래 기본법」 제2조제1호에 따른 전자문서를 말한다)
3. 개인신용정보의 제공 내용 및 제공 목적 등을 고려하여 정보 제공 동의의 안정성과 신뢰성이 확보될 수 있는 유무선 통신으로 개인비밀번호를 입력하는 방식
4. 유무선 통신으로 동의 내용을 해당 개인에게 알리고 동의를 받는 방법. 이 경우 본인 여부 및 동의 내용, 그에 대한 해당 개인의 답변을 음성녹음하는 등 증거자료를 확보·유지하여야 하며, 대통령령으로 정하는 바에 따른 사후 고지절차를 거친다.

5. 그 밖에 대통령령으로 정하는 방식

② 개인신용평가회사, 개인사업자신용평가회사, 기업신용조회회사 또는 신용정보집중기관으로부터 개인신용정보를 제공받으려는 자는 대통령령으로 정하는 바에 따라 해당 신용정보주체로부터 제1항 각 호의 어느 하나에 해당하는 방식으로 개인신용정보를 제공받을 때마다 개별적으로 동의(기존에 동의한 목적 또는 이용 범위에서 개인신용정보의 정확성·최신성을 유지하기 위한 경우는 제외한다)를 받아야 한다. 이 경우 개인신용정보를 제공받으려는 자는 개인신용정보의 조회 시 개인신용평점이 하락할 수 있는 때에는 해당 신용정보주체에게 이를 고지하여야 한다.

③ 개인신용평가회사, 개인사업자신용평가회사, 기업신용조회회사 또는 신용정보집중기관이 개인신용정보를 제2항에 따라 제공하는 경우에는 해당 개인신용정보를 제공받으려는 자가 제2항에 따른 동의를 받았는지를 대통령령으로 정하는 바에 따라 확인하여야 한다.

④ 신용정보회사등은 개인신용정보의 제공 및 활용과 관련하여 동의를 받을 때에는 대통령령으로 정하는 바에 따라 서비스 제공을 위하여 필수적 동의사항과 그 밖의 선택적 동의사항을 구분하여 설명한 후 각각 동의를 받아야 한다. 이 경우 필수적 동의사항은 서비스 제공과의 관련성을 설명하여야 하며, 선택적 동의사항은 정보제공에 동의하지 아니할 수 있다는 사실을 고지하여야 한다.

⑤ 신용정보회사등은 신용정보주체가 선택적 동의사항에 동의하지 아니한다는 이유로 신용정보주체에게 서비스의 제공을 거부하여서는 아니 된다.

다. 보호대책의 준용

그 밖에 신용정보회사등이 마련해야 할 가명정보에 대한 보호조치는 「신용정보업감독규정」 [별표 3]의 신용정보의 기술적·관리적·물리적 보안대책을 준용한다. 가명정보 및 추가정보의 보호에 관하여 「신용정보업감독규정」 [별표 3]과 신용정보업감독규정 [별표 8]이 경합하는 때에는 [별표 8]을 우선 적용한다.

Ⅲ. 익명처리 및 적정성 평가

1. 개요

가. 익명처리

신용정보회사등은 신용정보법에 따라 더 이상 특정 개인인 신용정보주체를 알아볼 수 없도록 개인신용정보를 익명처리한 후 이를 회사 내부에서 이용하거나 제3자에게 제공할 수 있다.

나. 익명처리 단계

1) 익명처리

정보집합물(데이터셋)에서 개인을 식별할 수 있는 요소를 전부 또는 일부 삭제하거나 대체하는 등의 방법을 활용하여 더 이상 특정 개인인 신용정보주체를 알아볼 수 없도록 조치한다.

2) 적정성 평가

다른 정보와 결합하여 더 이상 특정 개인인 신용정보주체를 알아볼 수 없도록 적절하게 익명처리하였는지 평가한다. 신용정보회사등은 금융위원회에 익명처리 적정성 심사를 요청할 수 있다.

다. 기록보존의 의무

신용정보회사등은 개인신용정보를 익명처리를 한 경우에는 다음의 조치 기록을 3년간 보존하여야 한다.

1) 익명처리한 날짜
2) 익명정보의 항목
3) 익명처리한 사유와 근거

2. 익명처리 방법

가. 속성 분류 및 익명처리 적용 기준

1) 신용정보회사등은 익명처리의 대상이 되는 정보를 식별자, 개인식별가능정보로 분류한 후 적정한 익명처리 기법을 적용하여야 한다.

※ 익명처리 대상 정보의 분류는 익명처리 목적 및 이용·제공 환경 등에 따라 달라질 수 있음

2) 익명처리시 식별자는 삭제하여야 하며, 부득이하게 정보 이용 목적상 필요한 경우에는 적절하게 익명처리를 한 후 이용하여야 한다.

3) 개인식별가능정보 중 개인 식별 가능성이 높은 속성은 그 정도에 맞추어 익명처리 수준을 높이는 등의 조치를 취하여야 한다.

※ k-익명성 모델 등의 기법도 적용 가능('붙임 1'의 '9. 프라이버시 보호 모델' 참조)

4) 개인식별가능정보 중 개인 식별 가능성이 낮은 속성은 활용 목적, 해당 정보의 특성, 다른 정보와의 결합 등을 고려하여 필요시 동질성 공격, 배경지식 공격 등의 다양한 위험을 제거하기 위하여 추가로 익명처리 기법을 적용하여야 한다.

※ '붙임 1'의 '9. 프라이버시 보호 모델' 참조

나. 익명처리 기법

1) 총계처리, 데이터 삭제, 데이터 범주화, 데이터 마스킹 및 프라이버시 보호모델 등 여러 가지 기법을 단독 또는 복합적으로 활용해야 한다.

2) 각각의 기법에는 이를 구현할 수 있는 다양한 세부기술이 있으며, 데이터 이용 목적과 기법별 장·단점 등을 고려하여 적절한 기법 및 세부기술을 선택하고 활용해야 한다('붙임 1' 참조).

< 활용 가능한 가명·익명처리 기술(세부 내용은 '붙임 1' 참조) >

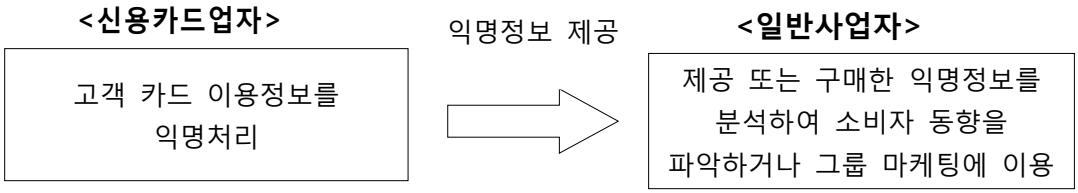
◎ 신용정보회사등은 익명처리시 본 안내서에서 설명하는 기술이라고 하더라도 해당 익명처리에 적합한 기술을 선별하여 적용하여야 한다. 또한 본 안내서에서 설명하지 않은 다른 기술이 더 적합하다고 판단될 경우에는 그것을 적용하여 익명처리를 할 수 있다.

구분		특징
통계 도구	표본추출	정보주체 별로 전체 모집단이 아닌 일부를 추출하여 사용하는 방법
	총계처리	속성값의 평균 또는 합계 등으로 처리
암호화 도구	결정적 암호화	동일한 키를 사용한 암호화 방식
	순서보존암호화	동일한 키로 암호화된 두 값이 암호문에서 같은 순서를 유지
	형태보존암호화	원본 데이터와 같은 형식, 길이를 갖는 일련의 기호 형식으로 데이터 변환
	동형 암호화	복호화를 하지 않고 암호화된 상태로 덧셈, 뺄셈 등 연산 수행
	동형 비밀분산	데이터 레코드 내에 식별자 또는 민감속성자를 k개의 분산 비밀정보값으로 대체
삭제 기법	마스킹	특정 속성값을 '**' 또는 'OO' 등으로 대체
	로컬 삭제	특정 속성값을 해당 레코드에서 삭제(부분 삭제)
	레코드 삭제	데이터에서 특이치(outlier) 등 특별히 구분되는 속성값을 포함하고 있는 레코드를 제거
가명화 기법	-	정보주체의 식별자를 각 정보주체에 대해 특별 생성된 간접식별자로 대체하는 기법
해부화	-	기존 하나의 데이터셋(테이블)을 2개의 데이터셋으로 분리하는 방식

구분		특징
일반화 기법	라운딩	특정 기준값을 베이스로 올림 또는 반올림 처리
	상·하단 코딩	최댓값과 최솟값을 정하여 주어진 값을 최댓값 또는 최솟값으로 대체
	속성집합을 단일속성값으로 결합	범주화
	로컬 일반화	특이값이 포함된 집단에 대해서만 일반화를 적용하는 기법
무작위화 기법	순열	속성값을 수정하지 않고 레코드 간에 속성값을 재정렬(교환)
	잡음 추가	원본 속성의 통계적 특징을 최대한 유지하면서 해당 속성값에 무작위 값을 곱하거나 더하여 추가
	부분 총계	총계처리의 일종으로 연속 속성(즉, 동질집합 내 레코드들)의 모든 값을 특정 알고리즘으로 계산된 평균치로 대체
재현데이터	-	실제 데이터를 기반으로 원 데이터의 분포를 추정한 후 이를 바탕으로 통계적 및 확률적으로 원본 데이터와 유사한 가상의 데이터
프라이버시 보호 모델	k-익명성 모델	동일한 속성을 가지는 레코드가 최소한 k개 이상 존재하도록 하여 프라이버시를 보호
	l-다양성 모델	동질집합(equivalent class)의 민감속성정보(sensitive attribute)가 최소한 l개의 다양한 속성을 가지도록 하여 k-익명성의 취약점(동질성 공격, 배경지식 공격)을 보완함
	t-근접성 모델	특정 동질집합의 기타속성자 분포와 전체 데이터의 기타 속성자 분포 차이를 t 이하가 되도록 조정
	차분 프라이버시 보호 모델	1개의 레코드가 차이 나는 두 DB의 차이(확률 분포)를 기준으로 하는 프라이버시 모델

[익명처리 예시]

▷ 처리 목적 : 신용카드업자가 카드이용정보를 익명처리 후 일반사업자에게 제공하고 일반사업자는 이를 분석하여 마케팅 등에 이용하고자 함
▷ 식별자 처리 : 삭제 적용
▷ 개인식별 가능정보 처리 : k-익명성, 범주화, 일반화, 삭제 등 적용



가. 보유 정보 샘플

ID	성명	카드 번호	전화 번호	성별	생년 월일	주소	근무 처	연봉 (만원)	내부 신용 등급	연체 잔고 (만원)	결제 기관
19342	홍길동	3779 4593 3043 3921 3943	010 -3355 -0934	남	1972. 9.9.	서울시 강남구 역삼동 332-1	OO 자동차	4,500	5	35	국민 은행
19354	김철수	4832 2332 2344 4399	02 -531 -9834	남	1980. 4.16.	서울시 마포구 송내동 334-1	OO 은행	6,000	2	0	새마을 금고
19445	전지연	4523 3234 9843 0394	010 -9290 -3344	여	1979. 5.23.	경기도 광주시 송정동 786-1	OO 공사	5,500	5	125	우리 은행
20221	박식별	4932 3453 5943 4321	010 -2891 -3322	여	1983. 12.3.	경기도 용인시 죽전동 33-11	OO 법무법인	7,000	1	0	기업 은행
…	…	…	…	…	…	…	…	…	…	…	…

나. 목적, 위험성 등을 고려한 익명처리

○ 익명처리 목적 및 이용·제공 환경, 다른 정보와의 결합, 동질성 공격, 배경지식 공격 등의 다양한 위험을 고려하여 익명처리를 적용

※ 예시의 속성분류와 적용한 익명기술이 절대적인 것은 아님

구분	속성	위험	익명처리 기술
ID	식별자	신용카드사에서 개인을 식별하기 위한 ID이므로 개인이 특정될 가능성이 있음	삭제
성명		식별자로 개인을 특정할 수 있음	삭제
카드번호		식별자이며 다른 사업자도 보유하고 있을 수 있으므로 개인이 식별될 위험이 존재	삭제
전화번호		식별자이며 다른 사업자도 보유하고 있을 수 있으므로 개인이 식별될 위험이 존재	삭제
성별	개인식별 가능정보* (k-익명성 적용)	생년월일, 주소 등의 정보와 조합하여 개인이 식별 될 수 있음	생년월일, 주소 등의 정보를 익명처리하여 위험 해소
생년월일		주소, 성별 등의 정보와 조합하여 개인이 식별 될 수 있음	생일을 삭제하고, 나이를 연령대로 범주화(20대/30대 등)
주소		성별, 생년월일 등의 정보와 조합하여 개인이 식별 될 수 있음	동 이하 주소는 삭제, k-익명성 수준을 만족하지 못할 경우 구 이하 주소 삭제 등 프라이버시 보호 모델 수준을 만족하도록 조치
직업	개인식별 가능정보 (k-익명성 미적용)	다른 사업자도 보유하고 있을 수 있으며 다른 정보와 조합하여 개인이 식별 될 위험이 존재	자영업, 공무원, 회사원, 기타로 일반화
연봉		수입이 너무 많거나 적은 경우 다른 정보와 조합하여 개인이 식별될 위험이 존재	연봉의 분포를 고려하여 3단계로 범주화
내부신용등급		이미 등급화 된 정보이며 개인에 민감한 정보	내부신용등급을 5단계로 범주화
연체잔고		개인의 소득을 추정할 수 있는 민감한 정보	연체잔고의 분포를 고려하여 10단계로 범주화
결제기관		익명정보를 제공받을 사업자가 필요한 정보가 아님	삭제

* 개인식별 가능정보 중 k-익명성 적용 대상 정보는 다른정보와 결합시 개인식별 가능성 등을 고려하여 선정

3. 적정성 평가

가. 신용정보회사등은 신용정보법에 따라 개인신용정보에 대하여 익명처리가 적정하게 이루어졌는지 금융위원회에 적정성 심사를 요청할 수 있다. 금융위원회는 익명처리의 적정성 심사 및 익명처리의 적정성 인정업무를 데이터전문기관에 위탁한다.

※ 금융위원회가 적정하게 익명처리가 이루어졌다고 인정한 경우 더 이상 해당 개인인 신용정보주체를 알아볼 수 없는 정보로 추정(개인신용정보에 해당한다는 반증이 없는 한 개인신용정보가 아니지만 개인신용정보라는 반증이 나오는 경우 개인신용정보로 본다는 의미임)

◎「신용정보법」 제40조의2(가명처리·익명처리에 관한 행위규칙)

③ 신용정보회사등은 개인신용정보에 대한 익명처리가 적정하게 이루어졌는지 여부에 대하여 금융위원회에 그 심사를 요청할 수 있다.

④ 금융위원회가 제3항의 요청에 따라 심사하여 적정하게 익명처리가 이루어졌다고 인정한 경우 더 이상 해당 개인인 신용정보주체를 알아볼 수 없는 정보로 추정한다.

⑤ 금융위원회는 제3항의 심사 및 제4항의 인정 업무에 대해서는 대통령령으로 정하는 바에 따라 제26조의4에 따른 데이터전문기관에 위탁할 수 있다.

◎「신용정보법」 시행령 제37조(권한의 위임 또는 위탁)

⑤ 금융위원회는 법 제40조의2제3항의 익명처리의 적정성 심사 및 법 제40조의2제4항의 익명처리의 적정성 인정업무를 데이터전문기관에 위탁한다.

나. 신용정보회사등은 익명처리가 적정하게 이루어졌는지 자체적으로 그 평가를 수행할 수 있다.

1) 금융위원회의 적정성 평가

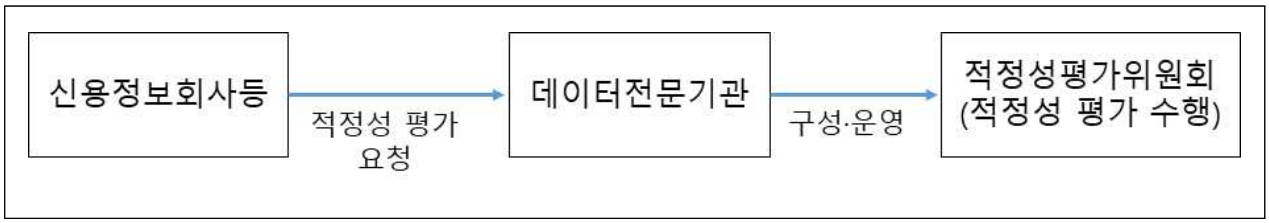
금융위원회는 익명처리 적정성 평가 관련 업무를 영제37조(권한의 위임 또는 위탁)에 따라 데이터전문기관에 위탁하였으며, 데이터전문기관은 결합정보에 대한 적정성 평가 및 익명처리 적정성 평가 업무를 수행하기 위하여 적정성평가위원회를 구성·운영할 수 있다. 데이터전문

기관은 가명처리·익명처리 관련 기법, 법률 등에 대한 전문지식 수준 등에 따른 자격 기준을 마련하여 적정성 평가위원회를 구성·운영할 수 있다.

< (예시) 적정성 평가위원 자격 기준 >

구분	자격 기준
법률전문가	1. 변호사의 자격을 소지한 자로서 1년 이상 관련 법률 업무(개인정보 보호, 데이터 가공·분석·활용, 데이터 가명·익명처리 및 적정성 평가 등 관련 법률 자문 또는 지원 업무로, 이하 동일)를 수행한 경력이 있는 자 2. 법학박사 학위를 취득한 자로서 2년 이상 관련 법률 업무를 수행한 경력이 있는 자 3. 법학석사 학위를 취득한 자로서 4년 이상 관련 법률 업무를 수행한 경력이 있는 자 4. 법학학사 학위를 취득한 자로서 6년 이상 관련 법률 업무를 수행한 경력이 있는 자 5. 8년 이상 관련 법률 업무를 수행한 경력이 있는 자
기술전문가	1. 국가기술자격법에 따른 정보관리기술사, 컴퓨터시스템응용기술사, 정보통신기술사 자격을 취득한 자로서 2년 이상 관련 업무(개인정보 보호, 데이터 가공·분석·활용, 데이터 가명·익명처리 및 적정성 평가 등으로, 이하 동일)를 수행한 경력이 있는 자 2. 관련 분야(컴퓨터공학, 정보보호학, 데이터베이스공학, 통계학, 수학 등으로, 이하 동일)에서 박사 학위를 취득한 자로서 2년 이상 관련 업무를 수행한 경력이 있는 자 3. 관련 분야에서 석사 학위를 취득한 자로서, 4년 이상 관련 업무를 수행한 경력이 있는 자 4. 관련 분야에서 학사 학위를 취득한 자로서 6년 이상 관련 업무를 수행한 경력이 있는 자 5. 관련 분야에서 8년 이상 관련 업무를 수행한 경력이 있는 자

< 적정성 평가 절차 >



적정성 평가를 신청할 경우, 신용정보회사등은 적정성평가위원회가 해당 익명처리의 적정성을 판단할 수 있도록 데이터 명세, 익명처리 현황 등을 포함한 기초자료와 자체 기준 및 절차에 따라 익명처리된 데이터를 제출하여야 한다(‘붙임 2’ 참조).

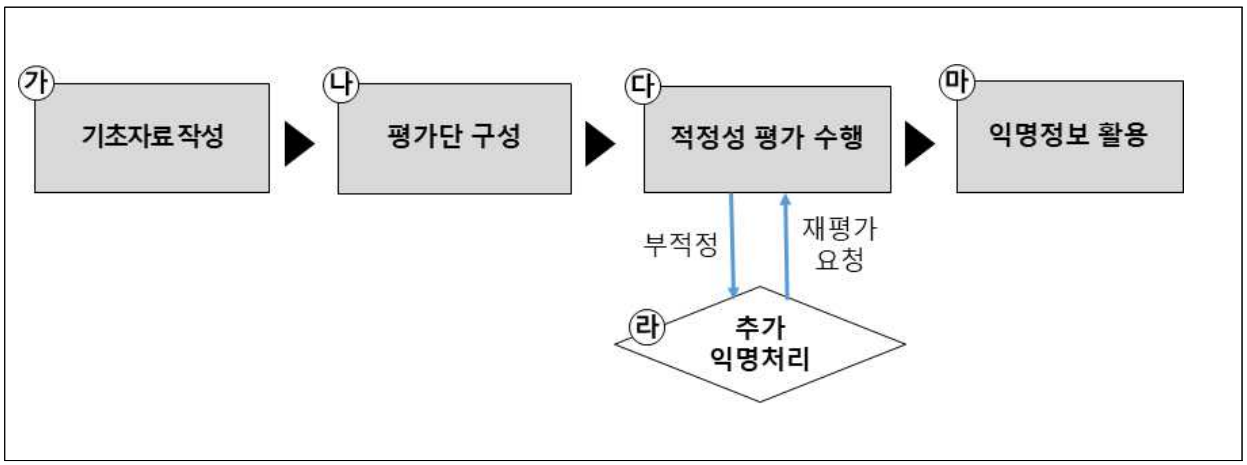
※ 제출해야 하는 기초자료의 내용은 데이터전문기관별로 상이할 수 있음

2) 자체 적정성 평가

신용정보회사등은 자체적으로 익명처리에 대한 적정성 평가를 수행할 수 있으며 다음은 이러한 평가수행 절차의 예시이다.

※ 해당 절차는 예시이며, 신용정보회사등은 자체 규정을 마련하여 수행 가능

< (예시) 자체 적정성 평가 수행 절차 >



가) **(기초자료 작성)** 신용정보회사등은 적정성 평가에 필요한 데이터 명세, 익명처리 현황, 이용기관의 관리 수준 등이 포함된 기초자료를 작성함

나) **(평가단 구성)** 신용정보관리·보호인이 전문가 3명 이상으로 적정성 평가단을 구성하되, 외부전문가가 평가단의 과반 이상이 되도록 함

※ 평가단 전문가는 '(예시) 적정성 평가위원 자격기준'(p.59)을 참고하여 구성

다) **(적정성 평가 수행)** 평가단은 신용정보회사등이 작성한 기초자료와 k-익명성 모델 등을 활용하여 익명처리 수준의 적정성을 평가함

라) **(추가 익명처리)** 신용정보회사등은 평가결과가 '부적정'인 경우 평가단의 의견을 반영하여 추가 익명처리를 수행한 후 그 적정성을 재평가 받음

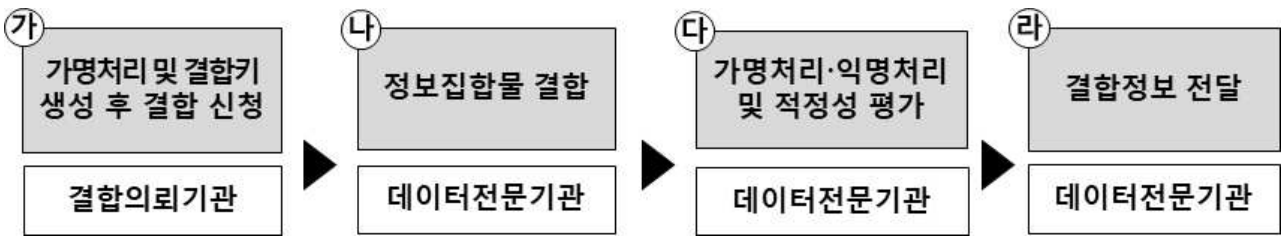
마) **(익명정보의 활용)** 익명처리가 적정하다고 평가받은 경우에는 해당 정보를 이용 또는 제공할 수 있음

Ⅳ. 정보집합물 결합

1. 개요

개정된 「신용정보법」은 신용정보회사등이 자신이 보유한 정보집합물을 제3자가 보유한 정보집합물을 데이터전문기관을 통하여 결합하는 것을 허용하고 있다.

< 정보집합물 결합 절차 개요 >



가. 가명처리 및 결합키 생성 후 정보집합물 결합 신청

결합의뢰기관은 결합 대상 정보집합물을 가명처리한 후 결합의뢰기관간 협의한 방식으로 결합키를 생성하고 데이터전문기관에 정보집합물 결합을 신청한다. 결합의뢰기관은 결합추진 여부를 결정하기 위하여 본 단계 전에 결합률 사전통지*를 데이터전문기관에 의뢰할 수 있다.

* 결합률 사전통지 : 결합의뢰기관간 협의된 방식으로 생성된 결합키를 데이터전문기관이 미리 전달받아, 이를 기준으로 매칭된 결합률을 결합의뢰기관에 통지

※ '결합률 사전통지'는 결합의뢰기관의 선택에 따른 절차로, 데이터전문기관에 따라 결합률 사전통지 서비스를 운영하지 않을 수 있음

나. 정보집합물 결합

결합의뢰기관은 데이터전문기관에 결합 신청이 접수되면, 결합 대상 정보집합물을 저장매체 또는 정보통신망을 통해 데이터전문기관이 정하는 안전한 방법으로 데이터전문기관에 전달한다. 데이터전문기관은 전달받은 복수의 정보집합물을 결합키를 기준으로 결합*한다.

* 데이터전문기관은 결합의뢰기관이 제출한 결합대상 정보집합물에서 결합키가 같은 레코드의 속성들을 결합하여 결합된 결과물만 양 결합의뢰기관에 전달함(결합되지 않은 결과물은 해당 정보집합물을 전달한 결합의뢰기관에만 전달 가능)

다. 가명·익명처리 및 적정성 평가

데이터전문기관은 결합정보를 결합의뢰기관의 선택에 따라* 가명처리 또는 익명처리를 추가로 수행한 후 적정성 평가를 진행한다. 적정성 평가 결과가 '적정'이 나올 때까지 가명처리 또는 익명처리를 수행한다.

* 결합의뢰기관이 가명정보를 요청한 경우에는 가명처리 및 가명처리 적정성 평가를, 익명정보를 요청한 경우에는 익명처리 및 익명처리 적정성 평가를 수행

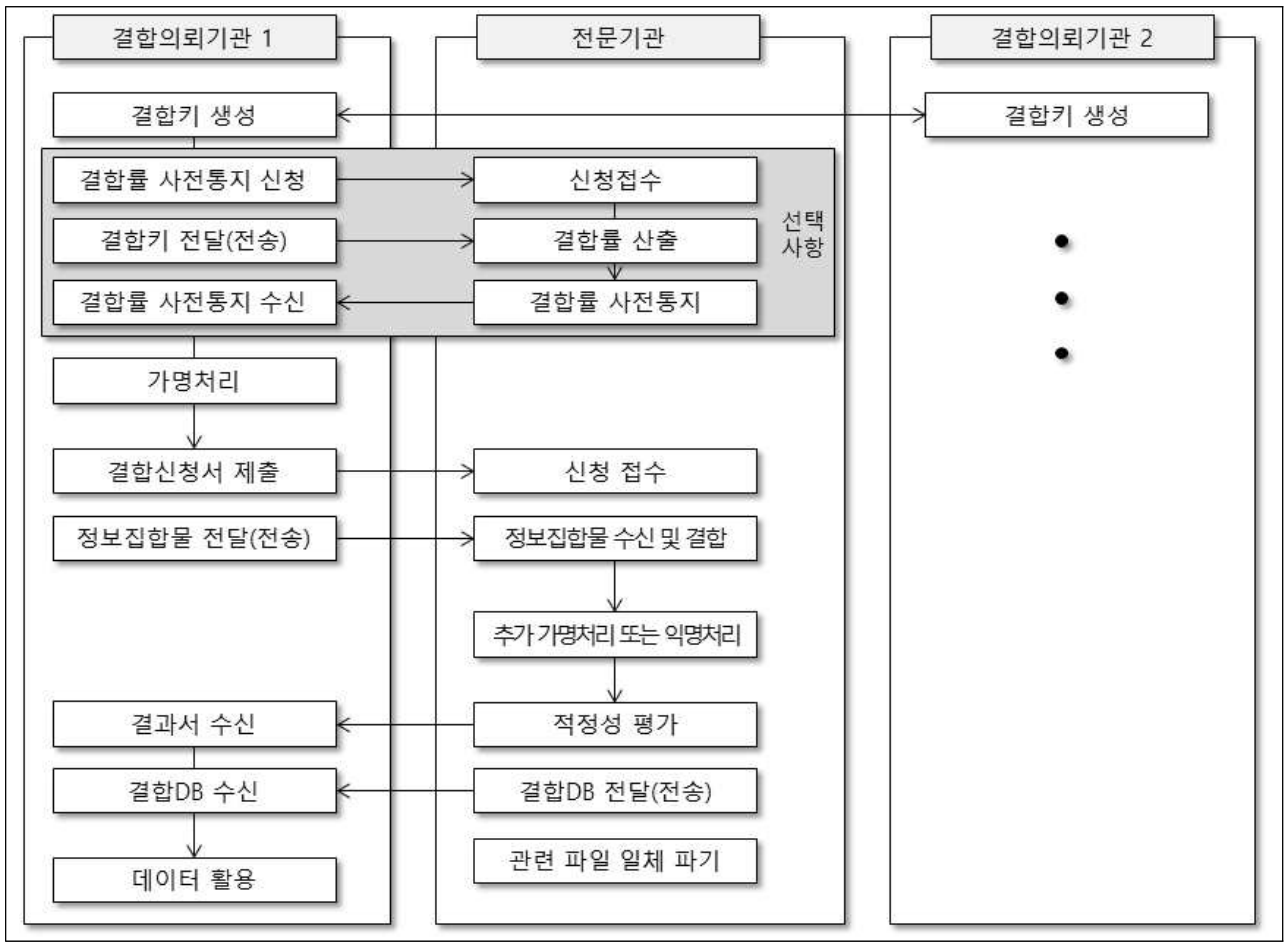
※ 데이터전문기관은 신청단계에서 결합의뢰기관의 선택에 따라 결합된 정보집합물을 가명처리 또는 익명처리한 후 전달(데이터전문기관이 분석공간을 제공할 경우, 가명·익명처리 적정성 평가가 완료된 데이터만 분석할 수 있음. 단, 결합정보를 반출하지 않는 조건 하에서는 가명처리 수준은 다를 수 있음)

라. 결합정보 전달

데이터전문기관은 적정성 평가까지 완료한 결합정보를 결합의뢰기관에 안전한 방법으로 전달(전송)한다. 결합정보가 정상적으로 수신되면 데이터전문기관은 관련 파일 일체를 지체 없이 파기한다.

2. 결합 절차

< 결합 세부 절차 >



※ 결합의뢰기관은 결합을 희망하는 상대 기관과 사전에 협의하여 결합 대상 정보집합물의 정보, 이용목적, 결합키 생성방법 등을 구체적으로 결정하여야 함

가. 결합키 생성

결합의뢰기관은 결합 상대기관과 협의한 방식으로 결합키를 생성한 후 결합 대상 데이터에 추가한다. 결합의뢰기관은 서로 협의하여 결합키의 생성 알고리즘과 생성시 활용할 입력정보(식별자)를 선택한다. 이 때, 결합키 생성을 위한 입력정보로 주민등록번호는 사용할 수 없으며, 생성된 결합키는 데이터의 신용정보주체를 유일하게 식별할 수 있는 값이어야 한다.

결합의뢰기관은 결합키가 포함된 정보집합물을 전문기관에 안전하게 전달해야 하며, 일방향 해시함수 등 결합키 생성에 관한 정보는 데이터전문기관과 공유할 수 없다. 결합의뢰기관은 생성한 결합키를 상호 공유할 수 없다. 결합키 생성을 위한 입력정보로 식별자에 해당하는 CI(Connecting Information)값을 사용하는 경우, 전체 CI값을 사용할 수 없고 재식별 위험이 없는 범위 내에서 CI 값의 일부를 사용할 수 있다.

[결합키 생성 방법 예시]

1. 결합키 생성절차

1) 결합의뢰기관은 결합키 생성을 위한 입력정보 및 인코딩 방식을 상호 협의하여 결정하여야 한다.

① 결합의뢰기관 상호 공통으로 보유하고 있는 정보를 활용하여 결합키 생성에 사용될 입력정보를 결정하여야 한다.

※ 예시) 성명 + 휴대폰번호

ㅇ 정보집합물 결합을 위한 결합키로 주민등록번호를 사용할 수 없으며, 결합키를 생성하기 위한 입력정보로도 활용할 수 없다(「개인정보 보호법」 제24조의2).

ㅇ 결합키 생성을 위해 입력되는 모든 정보는 비트(bit) 수준에서 동일한 방식으로 입력되어야 한다.

② 결합의뢰기관은 원본 정보와 결합시에도 개인을 식별할 수 없도록 ①에서 정한 결합키 입력정보에 솔트값을 추가하여 결합키를 생성하여야 한다.

※ 예시) 솔트값 : 'abcd1234'

③ 결합의뢰기관은 ①에서 정한 입력정보에 성명, 주소 등 한글과 같이 다국어가 포함될 경우 인코딩 방식이 동일해야 하므로 상호 동일한 인코딩(encoding)방식을 정하여 인코딩 한다.

※ 예시) utf-8, euc-kr 등

2) 결합의뢰기관은 결합키 생성 알고리즘 및 결합키 표현방식을 결정한다.

① 결합의뢰기관이 결합키를 생성할 알고리즘(일방향 해시함수 등)을 결정한다.

※ 예시) 일방향 해시함수(SHA256/384/512, HAS-160 등), XOR 등

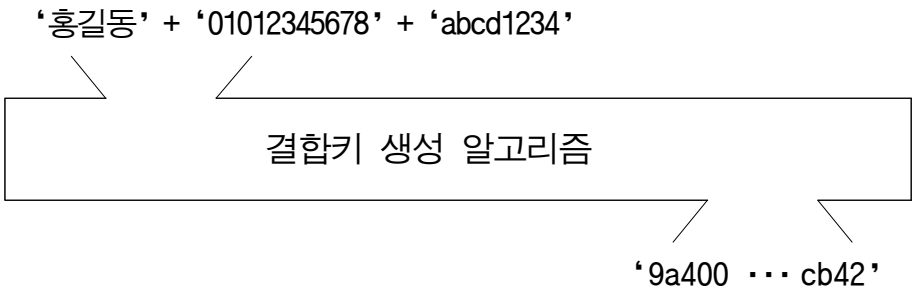
② 결합의뢰기관은 결정된 입력정보와 합의한 알고리즘으로 생성한 결합키의 표현방식을 결정한다.

※ 예시) base64, hexa 등

○ 결합의뢰기관은 입력정보의 예시를 활용하여 입력정보와 생성된 결합키를 상호 교환하여 일치 여부를 확인한다.

※ 결합의뢰기관은 결합키 생성방식의 오류 유무를 확인할 때 실제 데이터로 확인하여서는 아니 되며, 기관간 협의된 가짜 데이터를 활용(예 : 실존하지 않는 가상의 인물의 성명, 전화번호 등을 활용)

< 입력정보 예시 >



2. 일방향 해시함수를 이용한 결합키 생성 예시

1) 결합의뢰기관은 결합키 생성을 위한 입력정보 및 솔트값을 결정한다.

성명, 생년월일(6자리), 휴대폰번호, ‘abcd’

※ 결합의뢰기관 간 공통적으로 보유한 정보중 식별자 컬럼3개, 잡음으로 ‘abcd’를 선택

2) 결합의뢰기관은 입력정보에 대한 인코딩 방식을 결정한다.

UTF-8

※ 결합키와 정보집합물에 한글 등 다국어가 포함될 경우 결합의뢰기관 간 인코딩 방식을 정함

3) 결합의뢰기관은 결합키를 생성할 알고리즘(일방향 해시함수 등) 및 출력정보에 대한 인코딩 방식을 결정한다.

SHA256(해시함수), HEXA 소문자(출력값 인코딩)

※ 일방향 해시함수의 출력 정보는 이진(binary) 값이므로 텍스트 문자열로 변경하기 위해서는 결합의뢰기관 간 상호 협의하여 인코딩 방식을 정함

4) 결합의뢰기관은 예시를 통해 결합키를 생성하고 결합의뢰기관 간 상호 교환하여 결합키 일치 여부를 확인한다.

‘홍길동801212010123412 34abcd’

↓ UTF-8

SHA256()

↓ hexa

‘9a4005ebbdc5b5dcf399e1905c4291b48bdafb8549308eca84610b14f556cb42’

※ SHA256함수의 출력(32바이트)을 hexa로 표현하면 64바이트임

※ 결합률 사전통지는 결합의뢰기관이 선택적으로 활용하는 절차로서, 결합의뢰기관은 해당 절차를 거치지 않고 '⑤가명처리' 절차를 바로 진행할 수 있음(데이터전문기관에 따라서는 결합률 사전통지 제도를 운영하지 않을 수 있으므로 사전 확인 필요)

나. **(결합률 사전통지 신청)** 결합의뢰기관은 정보집합물 결합의 효과를 검토하기 위하여 정보집합물 결합률 사전통지 신청서를 작성하여 데이터전문기관에 제출한다.

※ 각 결합의뢰기관은 신청서를 개별 제출하여야 한다.

다. **(결합키 전달(전송))** 각 결합의뢰기관은 결합률 사전통지 신청이 접수되면 결합키를 데이터전문기관에 안전한 방법으로 전달(전송)한다.

라. **(결합률 사전통지 수신)** 데이터전문기관은 전달받은 결합키의 일치 여부를 확인하여 결합률을 산출하고 그 결과를 결합의뢰기관에게 통지한다.

마. **(가명처리)** 결합의뢰기관은 안내서에 따라 결합 대상 데이터의 가명처리를 수행한다.

바. **(신청서 제출)** 정보집합물 결합을 희망하는 복수의 결합의뢰기관은 정보집합물 결합 신청서를 데이터전문기관에 각각 제출하여야 한다.

○ 결합의뢰기관은 신청서를 제출할 때 결합 대상 정보집합물의 기초자료*를 첨부하여야 함

* 결합 대상 정보집합물의 정보(이름, 크기, 행과 열의 수 등) 및 컬럼별 정보(데이터 유형, 데이터 길이 등)가 포함되어야 함

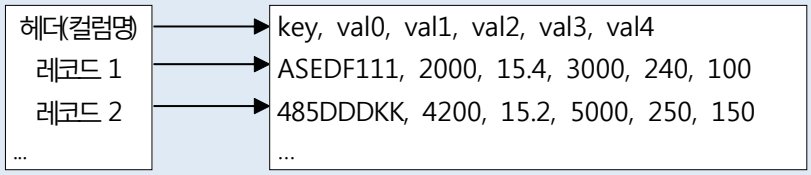
사. **(정보집합물 전달(전송))** 결합의뢰기관은 전문기관과 협의*하여 결합 대상 데이터를 저장매체 또는 정보통신망을 이용하여 데이터전문기관에 전달한다.

* 전문기관별로 정보집합물 전달방법이 상이할 수 있으므로 사전에 협의 필요

※ 결합의뢰기관은 정보집합물이 아래의 내용을 담은 CSV 형식이 되도록 하여야 함

▶ 헤더(컬럼명) + 레코드(결합키, 속성1, 속성2, 속성3, 속성4...)

< 예 시 >



아. **(정보집합물 결합)** 데이터전문기관은 정보집합물을 결합한 후 결합의뢰기관의 선택에 따라 가명처리 또는 익명처리를 수행하고 적정성 평가를 진행한다.

ㅇ 데이터전문기관은 가명·익명처리 목적, 가명정보 이용기관의 재식별 의도 및 능력, 가명정보 보호수준 및 신뢰도 분석 등을 고려하여 가명처리 또는 익명처리의 적정성을 평가

ㅇ 데이터전문기관은 결합 완료후 결합키를 삭제 또는 대체*

* 주기적·반복적 정보집합물 결합시 신용정보주체별로 연결해야 할 경우 연결키를 생성하여 결합키를 대체(본 안내서 'IV. 4. 외부기관간 주기적·반복적 정보집합물 결합 및 활용' 참고)

자. **(결합정보 전달(전송))** 데이터전문기관은 적정성 평가가 완료된 결합정보를 안전한 방법으로 전달(전송)한다.

ㅇ 결합정보가 정상적으로 수신되면 데이터전문기관은 관련 파일 일체를 파기

◎ 데이터전문기관의 분석 시스템 이용

▶ 데이터전문기관이 분석 시스템을 운영하는 경우, 결합의뢰기관은 데이터전문기관의 승인을 받아 분석 시스템에 보유 데이터를 반입하거나 분석결과를 반출할 수 있음

① **(분석공간 신청)** 결합의뢰기관은 결합정보를 데이터전문기관의 분석 시스템에서 분석하고자 할 경우, 결합정보 수신 전 데이터전문기관에 분석공간 이용을 신청

ㅇ **(보유 데이터 반입)** 결합의뢰기관이 보유 데이터와 결합정보를 함께 분석하고자 할 경우, 보유 데이터에 대한 반입을 신청

② **(결합정보 분석)** 결합의뢰기관은 결합정보 및 반입한 보유 데이터 등을 분석하여 이용목적을 충족하기 위한 분석결과를 도출

③ **(분석결과 반출)** 결합의뢰기관이 분석결과의 반출심사를 신청하고 데이터전문기관이 이를 승인하면 결합의뢰기관은 분석결과를 반출하여 이용목적에 맞게 활용 가능

④ **(관련 파일 일체 파기)** 결합의뢰기관의 분석결과 반출 후 데이터전문기관은 관련 파일 일체를 파기

⑤ **(반출입 기록 저장)** 데이터전문기관은 결합의뢰기관의 결합 관련 기록과 함께 결합의뢰기관의 분석공간 활용 및 데이터 반출입 기록을 저장하고 관리

차. **(결합정보 활용)** 결합의뢰기관은 결합된 정보집합물을 가명정보로 수신한 경우, 신청 단계에서 기재한 이용 목적에 한하여 활용하고 이에 대한 철저한 보호조치*를 수행하여야 한다.

※ 익명정보의 경우 목적 제한 없이 자유롭게 활용 가능하여 보호조치 불요

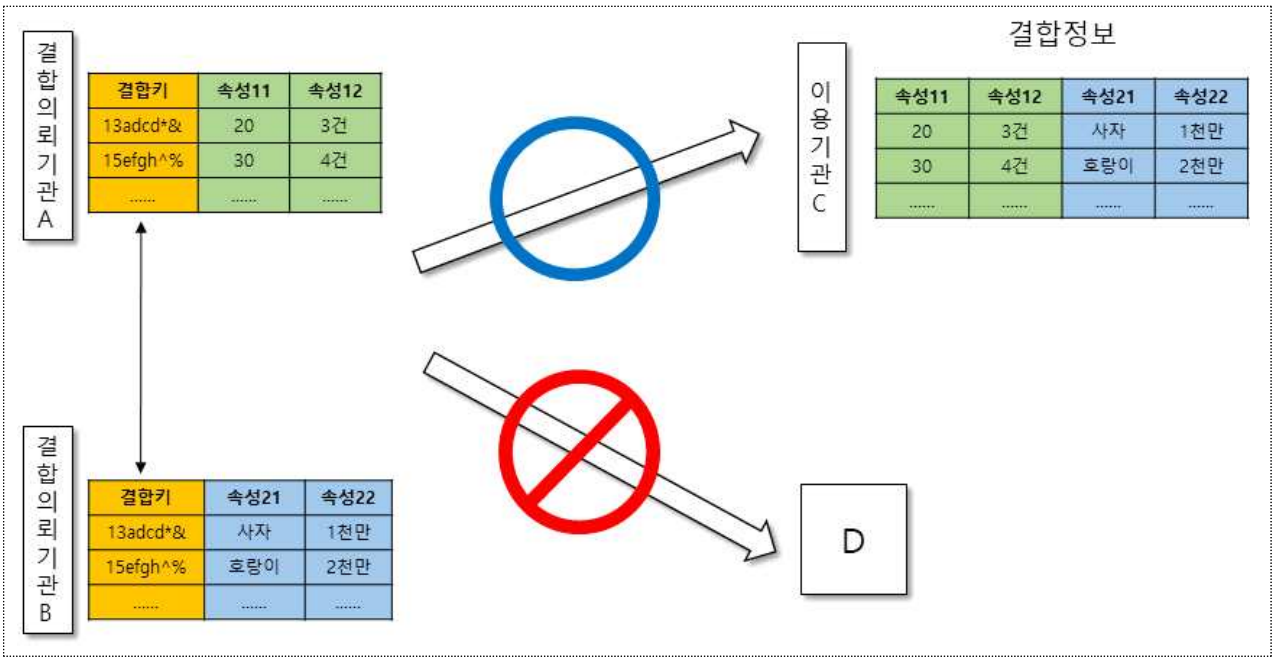
* 본 안내서 'Ⅱ. 4. 가명정보 및 추가정보에 관한 보호조치 기준(p36)' 참고

ㅇ 결합의뢰기관은 가명정보 보호조치 등 사후관리를 수행

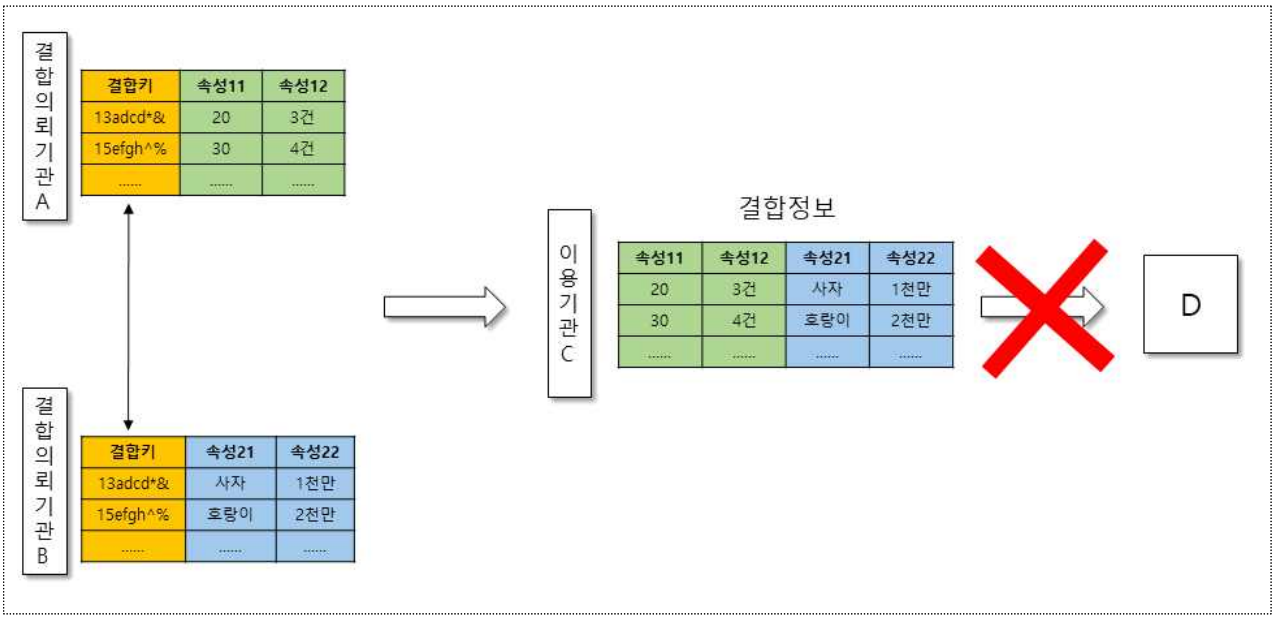
ㅇ 결합의뢰기관이 영리 또는 부정한 목적을 위하여 특정 개인을 알아볼 수 있게 결합정보 재결합하는 등 가명정보를 처리하는 것은 엄격히 금지됨(「신용정보법」 제40조의2제6항)

ㅇ 결합정보를 정보집합물 결합 신청시 결합정보를 이용할 기관으로 명시하지 않은 제3자에게 제공하는 것은 금지됨

※ (사례 1) 정보집합물 결합시 결합의뢰기관 A, B가 결합정보 이용기관을 C로 명시하였을 경우, 결합의뢰기관 A 또는 B가 추후 D에게 결합정보를 제공하는 것은 데이터전문기관이 해당 정보집합물 결합 및 적정성 평가시 고려하지 않은 사항으로서 허용되지 않음(결합의뢰기관 A, B가 데이터전문기관에 새로 정보집합물 결합을 신청하고 D의 가명정보 활용 목적, 보호수준 등을 고려한 적정성 평가를 수행 후 제공해야 함)



(사례 2) 정보집합물 결합시 결합의뢰기관 A, B가 결합정보를 이용기관 C에게 제공하였는데, 추후 C가 D에게 결합정보를 임의로 제공하는 것은 데이터전문기관이 해당 정보집합물 결합 및 적정성 평가시 고려하지 않은 사항으로서 허용되지 않음(결합의뢰기관 A, B가 데이터전문기관에 새로 정보집합물 결합을 신청하고 D의 가명정보 활용 목적, 보호수준 등을 고려한 적정성 평가를 수행 후 제공해야 함)



3. 데이터전문기관 보유 정보집합물과 외부 정보집합물과의 결합

전문기관이 보유한 정보집합물과 외부 기관의 정보집합물을 결합하는 경우 결합목적, 결합한 정보집합물 이용기관, 관련 대가 지급 여부 등을 종합적으로 고려하여 이해상충 발생가능성이 없어야 한다.

가. 결합목적

결합할 정보집합물을 보유한 데이터전문기관의 이익과 관련 여부 등을 보아 판단한다.

나. 이용기관

결합할 정보집합물을 보유한 데이터전문기관과 결합된 정보집합물을 이용하는 기관과의 연관성 등을 보아 판단*한다.

* 원칙적으로 데이터전문기관은 자기가 결합한 정보집합물의 이용기관이 될 수 없음 (원칙적으로 데이터전문기관이 결합된 정보집합물을 이용하고자 하는 경우에는 다른 데이터전문기관을 통하여 정보집합물 결합을 수행하여야 함)

다. 대가지급

데이터전문기관이 자체 보유 정보집합물과 외부기관의 정보집합물을 결합하여 결합된 정보집합물을 해당 외부기관에 전달시 자체 데이터 가공 및 결합 등 해당 업무처리에 소요된 실비 등의 범위 내에서 외부기관 등으로부터 대가를 받았는지 여부 등을 보아 판단한다.

4. 주기적·반복적 정보집합물 결합 및 활용

가. 개요

신용정보회사등이 제3의 기관과 정보집합물 결합을 추진할 때 추후 동일한 상대기관, 동일한 활용목적, 동일한 형태의 정보집합물을 주기적·반복적으로 결합할 필요가 있는 경우*에 해당한다. 이에 해당하는 경우, 신용정보회사등은 데이터전문기관에 정보집합물 결합을 의뢰할 때 주기적·반복적 정보집합물 결합 신청서를 함께 제출하여야 한다.

* 시계열 분석, 장기적 연구, 주기적 통계처리 등

< 주기적·반복적 정보집합물 결합 관련 고려사항 >

- ㅇ 신용정보주체의 추가, 삭제 등의 변경이 없어야 함
- ㅇ 데이터 속성(컬럼 구성)에 변화가 없어야 함 (속성 추가·삭제는 안됨)
- ㅇ 주기적·반복적 정보집합물 결합을 통해 시간순의 데이터 추가 등이 가능함
- ㅇ 데이터전문기관은 결합 후 해당 데이터를 삭제하고 연결키, 연결키 생성 알고리즘, 솔트값 등을 보유

< 주기적·반복적 정보집합물 결합 신청시 고려사항 >

조건 항목	조건 작성 시 고려사항
활용 목적	- 최초 결합과 동일한 목적으로 활용하여야 함
정보 구조	- 최초 결합데이터와 동일한 정보 구조를 유지하여야 함
이용 환경	- 최초 결합데이터와 동일한 이용자가 활용하여야 함
주기적·반복적 결합 기간	- 주기적·반복적 결합에 대한 기한 명시 필요(기한이 완료되면 데이터전문기관은 연결키 생성 알고리즘, 솔트값 등을 파기)

※ 본 고려사항은 최소요건이며, 주기적·반복적 정보집합물 결합 신청시 정보집합물 결합 신청서와 함께 작성하여 제출해야 함

주기적·반복적 정보집합물 결합시 결합정보를 신용정보주체별로 연결해야 할 필요가 있는 경우 데이터전문기관은 연결키를 생성하여 결합키를 대체한다.* 데이터전문기관은 적정성 평가 완료 후 결합정보를 결합의뢰기관에 전달하고 연결키 생성정보**를 이용한 경우에는 그 정보를 분리 보관하여야 한다.

* 신용정보주체별로 연결될 필요가 없는 경우에, 데이터전문기관은 결합키를 대체하지 않고 삭제

** 연결키, 연결키 생성 알고리즘, 솔트값 등

결합의뢰기관은 전달받은 결합정보를 데이터 특성*에 맞게 관리하고, 가명정보일 경우 연결키를 기준으로 연결하여 활용할 수 있다. 결합의뢰기관의 주기적·반복적 정보집합물 결합이 종료되면, 데이터전문기관은 연결키 생성정보를 파기하여야 한다.

* 'Ⅱ. 5. 가명정보 및 추가정보에 관한 보호조치 기준' 참고

나. 주기적·반복적 정보집합물 결합 절차

1) 최초 정보집합물 결합

결합의뢰기관은 정보집합물 결합 신청시 주기적·반복적 정보집합물 결합을 함께 신청한다. 데이터전문기관은 정보집합물 결합 후 결합키를 삭제 또는 대체*하고 적정성 평가를 완료한 후 결합정보를 결합의뢰기관에게 전송한다.

* 주기적·반복적 정보집합물 결합시 신용정보주체별로 연결해야 할 경우 연결키를 생성하여 결합키를 대체

2) 이후 주기적·반복적 정보집합물 결합

결합의뢰기관은 정보집합물 결합 신청시 주기적·반복적 정보집합물 결합을 함께 신청한다. 최초 결합시 결합의뢰기관이 결합키를 대체한 경우, 두 번째 결합부터는 최초 정보집합물 결합 때 저장했던 연결키 생성정보를 이용하여 연결키를 생성하고 결합키를 대체한다. 데이터전문기관은 적정성 평가를 완료한 후 결합정보를 결합의뢰기관에게 전송한다.

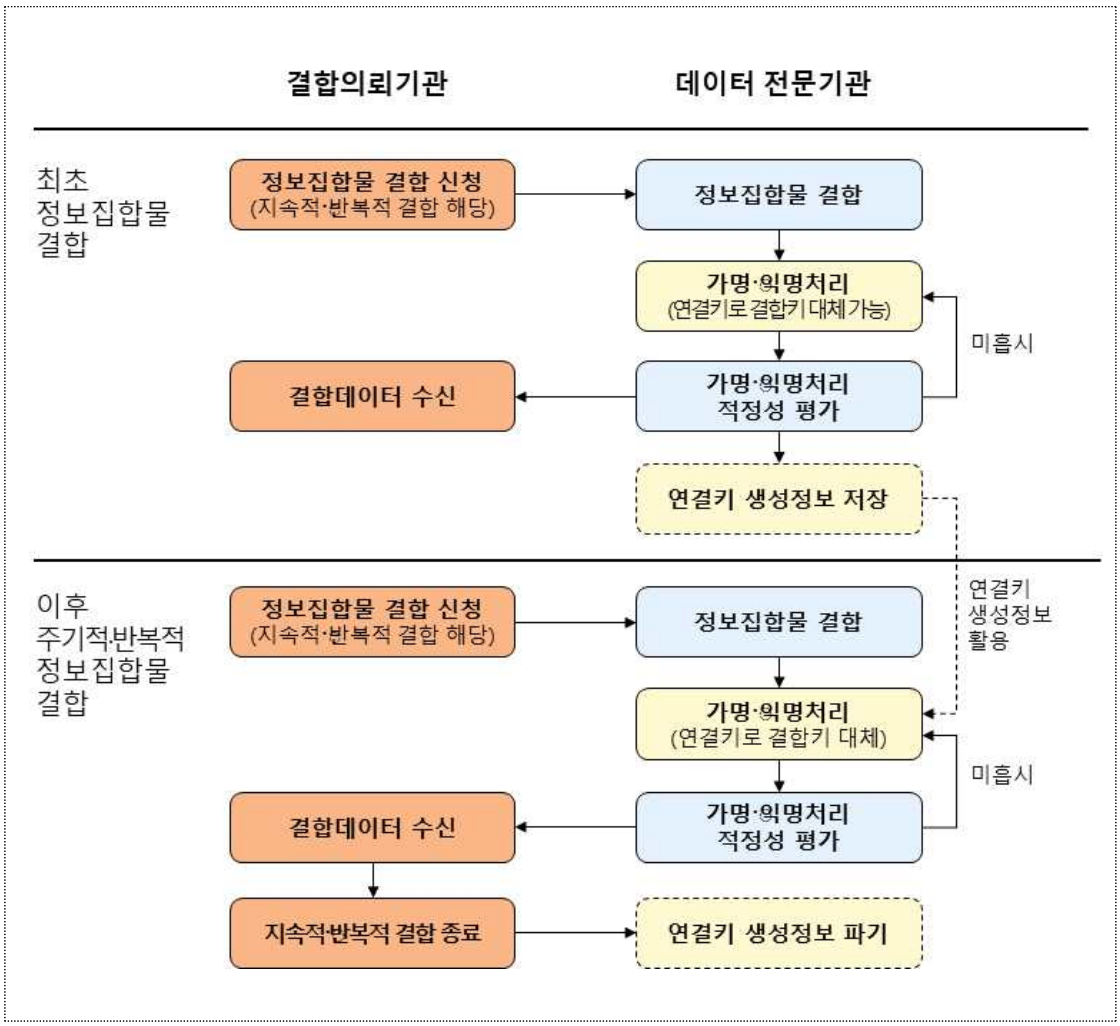
3) 주기적·반복적 정보집합물 결합 완료시

결합의뢰기관은 주기적·반복적 정보집합물 결합이 종료될 경우, 전문기관에 통지*하여야 한다. 주기적·반복적 정보집합물 결합이 종료되면 데이터전문기관은 해당 주기적·반복적 결합 건에 대한 연결키 생성정보 등을 파기**하여야 한다.

* 마지막 주기적·반복적 정보집합물 결합 신청시 신청기간에 '종료예정일' 입력

** 결합키를 대체한 경우에 해당

< 주기적·반복적 정보집합물 결합 절차 >



붙임 1. 가명·익명처리 기법

◎ ISO/IEC 20889, 『2019 개인정보 비식별 기술 가이드라인』(과학기술정보통신부.한국정보화진흥원, 2019.12.), 「한국신용정보원 가명처리·익명처리 전문가 양성」 연수자료 등을 참조

1. 통계 도구(Statistics tools)

데이터의 전체 구조를 변경하는 통계적 성질을 가진 기법으로, 일반적으로 가명처리·익명처리하거나 기법의 효율성을 높이기 위해 사용한다.

가. 표본추출(Sampling)

○ 정보주체 별로 전체 모집단이 아닌 일부를 추출하여 사용하는 방법

1) 확률적 표본추출

○ 무작위 표본추출, 계통적 표본추출, 층화 표본추출(모집단을 동질적인 몇 개의 층으로 나누고 각 층으로부터 무작위 표준을 추출하는 방식), 집락 표본추출 등

2) 비확률적 표본추출

○ 임의 표본 추출, 판단 표본추출, 할당 표본추출, 누적 표본추출 등

나. 총계처리(Aggregation)

○ 속성값의 평균 또는 합계 등으로 처리

○ 데이터 분석 목적에 따라 평균값, 최댓값, 최빈값, 중간값 중 어떤 값을 사용할지가 중요함

2. 암호화 도구(Cryptographic tools)

가명처리·익명처리 기법의 효율을 향상시키는 보안대책을 시행하기 위해, 또는 가명·익명처리 도구의 일부로 사용한다.

가. 결정적 암호화(Deterministic encryption)

○ 동일한 키를 사용한 암호화 방식(동일 속성값에 대한 암호화 결과도 동일)

○ 프라이버시 보존형 데이터마이닝(Privacy Preservation Data Mining, PPDM)*에 대해 데이터 유용성 보장

* 개인정보를 노출시키지 않으면서 데이터에 포함된 유용한 정보, 패턴 등을 찾아내는 데이터 마이닝 기법으로, 제3자가 데이터에 존재하는 개인을 인식하는 것을 방지하기 위하여 통계적 정보를 변경하는 원칙(일반화, 잡음 추가 등)

나. 순서보존암호화(Order-preserving encryption)

○ 동일한 키로 암호화된 두 값이 암호문에서 같은 순서를 유지

○ 결정성 암호화보다 더 높은 수준의 유용성을 제공

다. 형태보존암호화(Format-preserving encryption)

○ 원본 데이터와 같은 형식, 길이를 갖는 일련의 기호 형식으로 데이터 변환

○ 주민등록번호를 암호화 시 13자리 숫자로 암호화됨

라. 동형암호화(Homomorphic encryption)

○ 복호화를 하지 않고 암호화된 상태로 덧셈, 뺄셈 등 연산 수행 가능

○ 데이터 유용성을 보장하는 것이 장점이지만, 결정성 암호화에 비해 성능이 낮고 높은 저장 비용이 발생하는 단점이 존재함

마. 동형비밀분산(Homomorphic secret sharing)

○ 데이터 레코드 내에 식별자 또는 민감한 정보를 k개의 분산 비밀정보 값으로 대체(레코드 내의 민감속성자를 k명의 소유자가 나눠 갖는 개념)

○ 통제된 재식별이 필요할 경우 분산 비밀정보를 보유한 k명의 소유자가 모두 동의할 때 가능함

3. 삭제 기법(Suppression techniques)

레코드, 속성값, 또는 데이터셋에서 선택된 레코드를 제거하는 기법이다.

가. 마스킹(Masking)

○ 특정 속성값을 '**' 또는 'OO' 등으로 대체

나. 로컬 삭제(Local suppression)

○ 특정 속성값을 해당 레코드에서 삭제(부분 삭제)

○ 일반적으로 '일반화'가 적용된 이후에도 여전히 존재하는 개인식별가능정보들의 희귀한 값들을 제거하는 데 주로 사용됨

다. 레코드 삭제(Record Suppression)

○ 데이터에서 특이치(outlier) 등 특별히 구분되는 속성값을 포함하고 있는 레코드를 제거

4. 가명화 기법(Pseudonymization techniques)

◎ 개정 신용정보법의 '가명처리'와는 구별되는 개념으로, ISO/IEC 20889의 가명화 기법은 원본 데이터의 식별자를 다른 값으로 치환하는 것에 한정되는 반면, 개정 「신용정보법」의 가명처리는 식별자 치환 뿐만 아니라 필요시 다른 속성자를 삭제하거나 라운딩하는 등의 기법도 추가로 적용하는 것을 포함

ㅇ 정보주체의 식별자를 각 정보주체에 대해 특별 생성된 대체값으로 대체하는 기법으로, 정보주체의 신원을 노출하지 않고 다른 데이터셋의 관련 레코드와 연결할 수 있도록 함

ㅇ 가명화만을 단독으로 사용하는 경우, 특정(single out)되는 리스크를 줄일 수는 없음

ㅇ 일반적으로 통제된 재식별 프로세스에 사용될 수 있는 매핑테이블, 암호키 등과 같은 추가정보가 생성됨(단, 추가정보는 적절한 기술적·관리적 조치에 의해 보호되어야 함)

- 양방향 암호화, 일방향 암호화, 토큰 기법, 매핑테이블 등을 활용

5. 해부화(Anatomization)

기존 하나의 데이터셋(테이블)을 2개의 데이터셋으로 분리하는 방식으로서, 식별자 부분과 데이터(그외 속성자) 부분을 분리하는 기법이다. 하나의 값을 여러 개로 나누는 것도 포함한다(‘코드화’ 라고도 함).

※ 예시) 남(1,3,5,7,9), 여(2,4,6,8,10)

ㅇ 해부화는 데이터를 변경시키지 않고 구조만을 변경시키는 기법

6. 일반화 기법(Generalization techniques)

범주화로도 불리며, 특정한 값을 상위의 속성으로 대체하는 기법이다.

가. 라운딩(Rounding)

ㅇ 특정 기준값을 베이스로 올림 또는 반올림 처리

※ 랜덤라운딩(라운딩의 자릿수와 기준값을 자유롭게 지정), 제어라운딩

나. 상·하단 코딩(Top/Bottom coding)

○ 최댓값과 최솟값을 정하여 주어진 값을 최댓값 또는 최솟값으로 대체

○ 데이터 분포에서 상단 및/또는 하단 영역에 적용(나이 85세 이상 등)

다. 속성집합을 단일 속성값으로 결합(combining a set of attributes into a single attribute)

○ 범주화

라. 로컬 일반화(Local generalization)

○ 속성값 중에 특이값(outlier)이 존재할 경우 이를 제거하기 위해 적용

○ 특이값이 포함된 집단에 대해서만 일반화를 적용하는 기법

7. 무작위화 기법(Randomization techniques)

속성값을 수정하여 원래 값과 다르게 변형되도록 하는 기법으로, 추론시도의 효율성을 감소시키는 기법(perturbation)이라고도 한다.

가. 순열(Permutation)

○ 특정 컬럼의 데이터를 무작위로 순서를 변경(교환)

나. 잡음 추가(Noise addition)

○ 원본 속성의 통계적 특징을 최대한 유지하면서 해당 속성값에 무작위 값을 곱하거나 더하여 추가

○ 연관 컬럼이 존재하는 경우 동일한 정보의 잡음을 추가해야 분석결과에 영향을 주지 않음(분포, 평균, 분산, 표준편차, 공분산, 상관관계 등 고려)

○ 예를 들어, 시작일에 5일의 노이즈를 더하면 종료일에 대해서도 동일하게 5일의 노이즈를 더해야 전체 기간에 영향을 주지 않음

다. 부분 총계(Microaggregation)

○ 총계처리의 일종으로 연속 속성(즉, 동질집합 내 레코드들)의 모든 값을 특정 알고리즘으로 계산된 평균치로 대체

8. 재현데이터(Synthetic data)

실제 데이터를 기반으로 원 데이터의 분포를 추정한 후 이를 바탕으로 생성한 데이터로서 실제 데이터는 아니나 통계적 및 확률적으로 원본 데이터와 유사한 가상의 데이터를 의미한다. 속성 내 일부만을 적용한 부분 합성과 전체 데이터셋에 적용한 완전 합성으로 나눌 수 있는데, 재현데이터 생성 방법은 다양하며 차분 프라이버시 보호모델을 활용하여 합성데이터를 생성하기도 한다.

9. 프라이버시 보호 모델

프라이버시 보호 수준을 통계적 기법을 활용하여 정량적으로 나타내는 방식으로, 개인이 직접 식별되는 것뿐만 아니라 추론을 통해 식별되는 것도 방지하는 것을 목적으로 한다. 주로 익명처리의 적정성 평가 기준으로 활용된다.

가. k-익명성 모델

○ 동일한 속성을 가지는 레코드가 최소한 k개 이상 존재하도록 하여 프라이버시를 보호(k=3일 경우, 동일한 개인식별가능정보 중 식별가능성이 높은 정보를 가지는 사람이 3명 이상 존재하여 특정 개인 식별이 불가)

< k-익명성 모델 적용 전 >

[식별자 제거 데이터]

연령	성별	카드 결제금액
60	남	320,000
62	여	600,000
61	남	500,000
27	남	1,500,000
29	남	1,000,000
27	여	1,750,000
26	여	1,400,000
60	여	150,000
61	남	145,000
60	여	402,000
28	남	1,330,000
25	여	1,220,000

연결공격 (Linkage Attak)

1:1

[확보된 공개 데이터]

이름	연령	성별
김철수	60	남
이민아	62	여
안상태	61	남
김상우	27	남
한기범	29	남
장아름	27	여
양수지	26	여
김다래	60	여
윤영하	61	남
김순자	60	여
김민영	28	남
유슬아	25	여

< k개(3개) 이상의 레코드가 존재하는 동질집합을 구성하여 1:1 연결 방지 >

[k-익명성 적용 데이터]

동질집합 (Equivalent Class)

연령	성별	카드 결제금액
60대	남	320,000
60대	남	500,000
60대	남	145,000
60대	여	600,000
60대	여	150,000
60대	여	402,000
20대	남	1,500,000
20대	남	1,000,000
20대	남	1,330,000
20대	여	1,750,000
20대	여	1,400,000
20대	여	1,220,000

k : 1

[확보된 공개 데이터]

이름	연령	성별
김철수	60	남
이민아	62	여
안상태	61	남
김상우	27	남
한기범	29	남
장아름	27	여
양수지	26	여
김다래	60	여
윤영하	61	남
김순자	60	여
김민영	28	남
유슬아	25	여

○ k-익명성 모델의 취약점

- 동질성 공격(Homogeneity attack): k-익명성에 의해 레코드들이 범주화 되었더라도 일부 정보들이 모두 같은 값을 가질 수 있기 때문에 데이터 집합에서 동일한 정보를 이용하여 공격 대상의 정보를 알아내는 공격
- 배경지식에 의한 공격(Background knowledge attack): 주어진 데이터 이외의 공격자의 배경 지식을 통해 공격 대상의 민감한 정보를 알아내는 공격(예를 들어, 여자는 전립선염에 걸릴 수 없다는 배경지식을 활용하여 개인정보 추론)

◎ 원인

▶ 다양성의 부족(lack of diversity)

▶ 조치 시 정보의 다양성을 고려하지 않음(동일한 정보를 가진 레코드가 하나의 동질집합으로 구성될 경우 동질성 공격에 무방비)

▶ 강한 배경지식(strong background knowledge) : 의료, 금융, 교육 등 영역별 전문지식

나. l-다양성 모델

○ 동질집합(equivalent class)의 민감속성정보(sensitive attribute)가 최소한 l개의 다양한 속성을 가지도록 하여 k-익명성의 취약점(동질성 공격, 배경지식 공격)을 보완함

< l-다양성 모델 적용 전 >

[k-익명성 적용 데이터]

연령	성별	우편번호	신용등급
60대	남	180**	8
60대	남	180**	8
60대	남	180**	8
60대	남	180**	8
60대	여	180**	1
60대	여	180**	3
60대	여	180**	5
60대	여	180**	2

동질성 공격
(Homogeneity Attack)

ㅇㅇㅇ지역의 모든 60대 남자의 신용등급은 8등급

(추론 예) ㅇㅇㅇ지역에 사는 남자, 박철수의 신용등급은 8등급

동질집합 내에서 다양성이 부족하여 특정 개인의 정보 추론 가능

< l-다양성 모델 적용 후 >

동질집합이 l개(3개)의 다양한 민감정보(신용등급)을 가지도록 조정

[l-다양성 적용 데이터]

연령	성별	우편번호	신용등급
60대	*	1803*	8
60대	*	1803*	8
60대	*	1803*	5
60대	*	1803*	2
60대	*	1804*	1
60대	*	1804*	3
60대	*	1804*	8
60대	*	1804*	8

60대 남자의 신용등급 추론 가능성 낮아짐 (l값 = 3)

ㅇ l-다양성 모델의 취약점

- 쏠림 공격(skewness attack): 정보가 특정한 값에 쏠려 있는 경우 l-다양성 모델이 프라이버시를 보호하지 못함

◎ 쏠림 공격의 예

▶ 임의의 동질집합이 99개의 '위암 양성', 1개의 '위암 음성' 레코드로 구성되어 있다고 가정

▶ 공격자는 공격 대상이 99%의 확률로 '위암 양성'이라는 것을 알 수 있음

- 유사성 공격(similarity attack) : 익명처리된 레코드의 정보가 서로 비슷하다면 l-다양성 모델을 통해 처리되었다 할지라도 프라이버시가 노출될 수 있음

◎ 유사성 공격의 예

▶ 동질 집합의 병명이 서로 다르지만 의미가 유사할 수 있음(위궤양, 급성위염, 만성위염)

▶ 이를 통해, 공격자는 공격 대상의 질병이 '위'에 관련된 것이라는 사실을 알아낼 수 있음

다. t-근접성 모델

○ 특정 동질집합의 개인식별가능정보 분포와 전체 데이터의 개인식별가능정보 분포 차이를 t 이하가 되도록 조정(t가 0에 가까울수록 분포가 유사하며, 이를 통해 특정집단의 개인식별가능정보 추론문제 보완)

< t-근접성 모델 적용 전 >

[l-다양성 적용 데이터]

연령	성별	우편번호	소득
60대	남	180**	10,000
60대	남	180**	50,000
60대	남	180**	60,000
60대	남	180**	15,000
60대	여	180**	35,000,000
60대	여	180**	100,000,000
60대	여	180**	175,000,000
60대	여	180**	24,000,000

쏠림 공격 (Skewness Attack)

ㅇㅇㅇ지역의 60대 남자의 소득은 매우 낮다

(추론 예) ㅇㅇㅇ지역에 사는 남자, 박철수의 소득은 매우 낮다

특정한 값에 쏠린 특성을 이용하여 개인의 정보 추론 가능

< t-근접성 모델 적용 후 >

[t-근접성 적용 데이터]

연령	성별	우편번호	소득
60대	*	1803*	10,000
60대	*	1803*	50,000
60대	*	1803*	175,000,000
60대	*	1803*	24,000,000
60대	*	1804*	35,000,000
60대	*	1804*	100,000,000
60대	*	1804*	60,000
60대	*	1804*	15,000

기타속성자 분포의 특성을 이용한 추론 방지

60대 남자의 소득수준 추론 가능성 낮아짐

참고	임의의 동질 집합에서 민감한 분포 Pec, 전체 데이터에 대한 민감한 정보의 분포 Q - 모든 동일 집합에 대하여 Pec와 Q의 차이(D[Pec, Q]) t를 계산 - 분포의 차이가 t 이하여야 한다는 정의에 따라, 가장 큰 분포의 차이값이 전체 데이터의 t근접성을 대표함

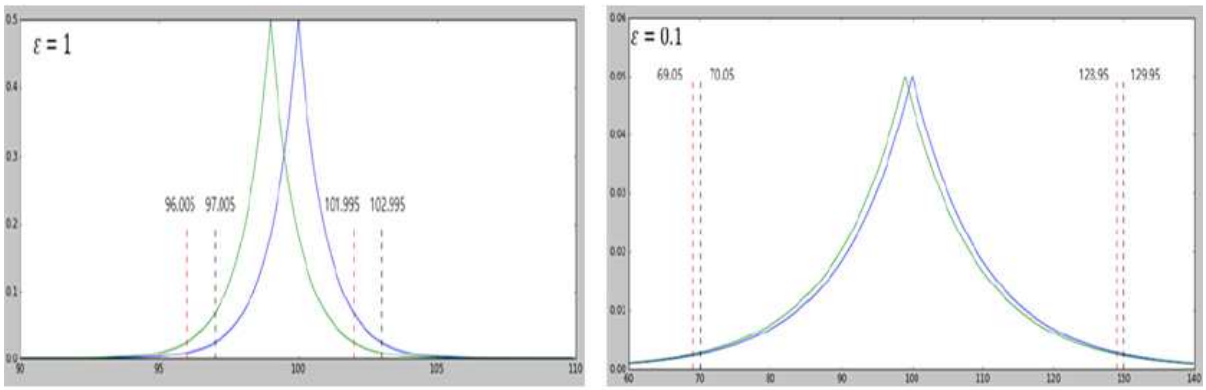
라. 차분 프라이버시 보호 모델(differential privacy)

○ k-익명성, l-다양성의 취약한 부분을 보완하기 위해 C.Dwork가 제안한 모형

○ 1개의 레코드가 차이 나는 두 DB의 차이(확률 분포)를 기준으로 하는 프라이버시 모델

- 두 DB간에 차이가 존재하면 차분 공격으로 알 수 있으나, 차이가 일정 크기 이하면 k>2 처럼 프라이버시 보호 수준이 생김
- 샘플링 또는 노이즈 추가 : 차이를 줄이기 위한 조치
- ε : 차이의 크기

$$P[K(D_1)\in S] \le \exp(\varepsilon)\times P[K(D_2)\in S] \text{ for all } S \subseteq Range(K)$$



Laplace(라플라스) 분포

원본

개인식별가능정보			신용불량여부
직업	성별	나이	
비-기술직	남자	[30~60)	Yes=2, No=17
전문직	남자	[30~60)	Yes=3, No=17
전문직	남자	[1~30)	Yes=1, No=20
전문직	여자	[30~60)	Yes=3, No=12
기술직	여자	[60~90)	Yes=2, No=23

[1/4 분기]

노이즈(샘플링 또는 가짜 레코드 삽입) 처리 후

개인식별가능정보			신용불량여부
직업	성별	나이	
비-기술직	남자	[30~60)	Yes=2+3, No=17+2
전문직	남자	[30~60)	Yes=3-1, No=17-2
전문직	남자	[1~30)	Yes=1+3, No=20+4
전문직	여자	[30~60)	Yes=3+5, No=12-3
기술직	여자	[60~90)	Yes=2+4, No=23+5

[1/4 분기]

개인식별가능정보			신용불량여부
직업	성별	나이	
비-기술직	남자	[30~60)	Yes=1, No=17
전문직	남자	[30~60)	Yes=3, No=17
전문직	남자	[1~30)	Yes=1, No=20
전문직	여자	[30~60)	Yes=3, No=12
기술직	여자	[60~90)	Yes=2, No=23

[2/4 분기]

개인식별가능정보			신용불량여부
직업	성별	나이	
비-기술직	남자	[30~60)	Yes=1+4, No=17+3
전문직	남자	[30~60)	Yes=3-2, No=17+1
전문직	남자	[1~30)	Yes=1+4, No=20+4
전문직	여자	[30~60)	Yes=3-1, No=12-2
기술직	여자	[60~90)	Yes=2+3, No=23+4

[2/4 분기]

출처: 공주대 최대선 교수

붙임 2. 익명처리 적정성 평가 기초자료 작성 방법(예시)

1. 기초자료 작성 항목

데이터 명세	■ 원본 데이터 크기(레코드 수)	필수
	■ 원본 데이터 생성 방법	필수
	■ 원본 데이터 관리 환경 (기술적, 물리적)	필수
	■ 원본 데이터 세부 항목별 명세 (범위, 개수 등)	필수
	■ 원본 예시(표)	필수
	■ (익명처리된)평가 대상 데이터 세부 항목별 명세	필수
	■ 평가 대상 데이터 (예시 또는 일부 레코드도 가능)	필수
익명처리 현황	■ 식별자, 개인식별가능정보 등 구분	필수
	■ 적용된 익명처리 기준(프라이버시 보호 모델 등) 및 수치	필수
	■ 익명처리 기법 · 세부기술	필수

※ 위의 항목에 대한 내용을 작성하여 익명처리 적정성 평가위원회에 제출

※ 분량 제한 없음

2. 기초자료 개요(작성예시 포함)

※ 모든 세부 항목에 대해서 아래의 개요에 맞게 작성하되 분량 제한 없음

가. 데이터 명세

1) 원본 데이터 특성

항목	내용
■ 데이터 크기(레코드 수)	*950 MB (레코드 수: 500만건)*
■ 데이터 생성 방법	*A카드 이용고객 정보에서 추출*
■ 데이터 관리 환경 (기술적, 물리적)	*접근통제, 계정관리, DB암호화 등이 적용된 DB 서버에 저장*

2) 원본 데이터 세부 항목별 명세

항목	상세 사항
이름	-
직업	*400가지(내부분류기준)*
성별	*남/녀*
고객등급	*1~9등급(내부분류기준)*
대출액 합계	*0~750,000,000(원)*

3) 원본 데이터 예시

이름	***직업***	***성별***	***고객등급***	***대출액 합계***
홍길동	*프로그래머*	*남*	*3*	*435,657,350*
이영애	*법정직 공무원*	*녀*	*8*	*126,450,000*

4) 평가 대상 데이터 세부 항목별 명세

항목	상세 사항
이름	*삭제*
직업	*30가지(범주.화)*
성별	*남→1, 여→2(가명화)*
고객등급	-
대출액 합계	*1,000,000(원) 범주로 구분*

5) 평가 대상 데이터 (예시, 일부 레코드)

직업	***성별***	***고객등급***	***대출액 합계***
IT종사자	*1*	*3*	*436,000,000*
공무원	*2*	*8*	*127,000,000*

나. 익명처리 현황

1) 익명처리 개요

항목	구분	익명처리 기법·기술	익명처리 기준 및 수치
이름	식별자	삭제	-
직업	개인식별가능정보	범주화	k-익명성(5)
성별	개인식별가능정보	가명화	
고객등급	개인식별가능정보	미적용	l-다양성(3)
대출액 합계	개인식별가능정보	라운딩, 범주화	

2) 익명처리 상세

항목 (조치기법)	원본 데이터	익명처리 데이터
직업 (범주화)	치과의사, 한의사	의사
	판사, 검사	법조인
	회사원, 공무원	급여소득자
	⋮	⋮
	주부, 학생	무직
성별 (가명화)	남자	1
	여자	2
대출액 합계 (라운딩 및 범주화)	1~1,000,000(원)	1,0.00,000(원)
	1,000,000~2,000,000(원)	2,000,000(원)
	⋮	⋮
	748,000,000~ 749,000,000(원)	749,000,000(원)
	749,000,000(원) 이상	750,000,000(원)

다. 개인식별가능정보 빈도수

	직업
구분	개인식별가능정보
학생	15,762
공무원	48,651
…	…
IT 종사자	6,489

	성별
구분	개인식별가능정보
1	3,349,574
2	1,650,426

	고객등급
구분	개인식별가능정보
1	267
2	3,496
…	…
9	596,748

	대출액 합계
구분	개인식별가능정보
0	614,756
1,000,000	3,695
…	…
1,000,000,000	24

참고문헌

○ 정부부처 합동, 『개인정보 비식별 조치 가이드라인』, 2016.6.

○ 과학기술정보통신부·한국정보화진흥원, 『2019 개인정보 비식별 기술 가이드라인』, 2019.12.

○ 한국금융연수원, 「한국신용정보원 가명처리・익명처리 전문가 양성」(연수자료), 2020.5.

○ 금융보안원, 「금융부문 암호기술 활용 가이드」(AGR-VII-2019-②-84), 2019.1.

○ ISO/IEC, “Privacy enhancing data de-identification terminology and classification of techniques”, ISO/IEC 20889, First edition, 2018.11.

○ ISO/IEC, “Health informatics — Pseudonymization”, ISO/IEC 25237, 2017.1.

○ ENISA, “Recommendations on shaping technology according to GDPR provisions: An overview on data pseudonymisation”, 2018.11.

○ Simson L. Garfinkel , “De-Identification of Personal Information”, NIST, NISTIR 8053, 2015.10.