# 가명정보 처리 가이드라인

2021. 10.

개인정보보호위원회
Personal Information Protection Commission

# 가명정보 처리 가이드라인

2021. 10.

개인정보보호위원회
Personal Information Protection Commission

본 가이드라인은 2021년 10월부터 12월까지
예시 및 Q&A 등이 추가 반영될 예정으로,
최신 버전의 가이드라인은 개인정보보호위원회
홈페이지(www.pipc.go.kr) 또는 가명정보
결합종합지원시스템(link.privacy.go.kr)에서
확인하시기 바랍니다.

# 목차

## 본편



## 부록

# 가명정보 처리 가이드라인

I

가이드라인
개요

II

## 가명처리

III

## 가명정보 결합 및 반출

IV

## 가명정보의 안전한 관리

# Ⅰ. 가이드라인 개요

## 1 목적

- 빅데이터, AI 등 다양한 융·복합 산업에서의 데이터 이용 수요가 급증하는 가운데, 데이터 활용의 핵심인 가명정보 활용을 위한 법적 근거가 마련됨에 따라,
- 가명정보 활용에 필요한 가명정보 처리 목적, 처리 절차 및 방법, 안전조치에 관한 사항 등을 안내하여 안전한 데이터 활용 환경을 마련하고자 함

○ 4차 산업혁명 시대 신성장 동력인 '데이터' 활용에 대한 시대적 요구를 반영한 데이터3법*이 시행('20.8.5.)되어 개인정보처리자가 통계작성, 과학적 연구, 공익적 기록보존 등을 위한 목적으로 개인정보를 가명처리하여 활용할 수 있는 기반이 새롭게 마련됨

* 개인정보 보호법, 정보통신망 이용촉진 및 정보보호 등에 관한 법률, 신용정보의 이용 및 보호에 관한 법률(이하 '신용정보법'이라 함)

○ 본 가이드라인은 「개인정보 보호법」(이하 '보호법'이라 함) 개정 및 시행('20.8.5.)으로 새롭게 도입된 '가명정보 처리에 관한 특례'(보호법 제3장 제3절)에 관한 설명과 구체적 사례를 제공함으로써 가명정보의 처리에 대한 이해를 돕고, 처리 과정에서 발생할 수 있는 개인정보 오·남용을 방지하여 안전한 가명정보 활용 방안을 안내하기 위해 작성하였음

※ 개인정보처리자가 법에 따른 규정을 준수한 경우 가이드라인 미준수를 사유로 처벌받지 않음
따라서, 개인정보처리자는 데이터의 관련 분야 및 특수성 등을 고려하여 상황에 따라 유동적으로 처리 가능

## 2 구성

- 본 가이드라인은 가명정보를 처리하는 실무자에게 도움이 되도록 본편과 부록으로 구성함.
  본편은 가명정보 처리에 대한 일반적인 사항들을, 부록은 가명처리와 관련된 각종 문서의 작성 방법 및 예시를 제공함

## 3 적용 대상

- 본 가이드라인의 적용 대상은 보호법에 근거한 가명정보 처리에 해당하지만, 관련 분야에서 개인정보위와 소관 부처가 공동으로 발간한 개인정보의 가명정보 처리에 관한 분야별 가이드라인*이 있는 경우에는 해당 분야의 가이드라인을 적용함

  * 금융분야 가명·익명처리 안내서, 보건의료 데이터 활용 가이드라인, 교육분야 가명·익명정보 처리 가이드라인, 공공분야 가명정보 제공 실무안내서 등

- 또한, 본 가이드라인은 '가명정보 처리에 관한 특례'(보호법 제3장 제3절)에 근거하여 통계작성, 과학적 연구, 공익적 기록보존 등을 위한 가명정보의 처리에 참고할 수 있도록 작성함

  ※ 보호법 제15조 제3항 및 제17조 제4항 등에 근거한 가명처리는 본 가이드라인의 적용대상이 아니지만, 가명처리에 관한 기술적 내용 등은 참고할 수 있음

**제15조(개인정보의 수집·이용)** 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.

**제17조(개인정보의 제공)** 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.

☑ 보호법 개정 및 시행('20.8.5.)으로, 「개인정보 비식별조치 가이드라인」('16)은 더 이상 현행법에 따른 가이드라인이 아니므로 활용하지 않음

# 4 용어 정리

| 구분 | 용어설명 |
|---|---|
| 개인정보 | 살아있는 개인에 관한 정보로서 다음의 정보를 포함함<br>- 성명, 주민등록번호 및 영상 등을 통하여 개인을 알아볼 수 있는 정보<br>- 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 함)<br>- 가명처리를 거쳐 생성된 정보로서 그 자체로는 특정 개인을 알아볼 수 없도록 처리한 정보 (이하 '가명정보'라 함)<br>※ 개인정보에 대한 판단기준은 개인정보처리자가 보유한 정보 또는 접근 가능한 권한 등 개인정보 처리 상황에 따라 다르게 판단되어야 함 |
| 개인정보파일 | 개인정보를 쉽게 검색할 수 있도록 일정한 규칙에 따라 체계적으로 배열하거나 구성한 개인정보의 집합물 |
| 익명정보 | 시간·비용·기술 등을 합리적으로 고려할 때 다른 정보를 사용하여도 더 이상 개인을 알아볼 수 없는 정보 |
| 가명처리 | 개인정보의 일부를 삭제하거나 일부 또는 전부를 대체하는 등의 방법으로 추가 정보(이하 '추가정보'라 함)가 없이는 특정 개인을 알아볼 수 없도록 처리하는 것 |
| 추가정보 | 개인정보의 전부 또는 일부를 대체하는 가명처리 과정에서 생성 또는 사용된 정보로서 특정 개인을 알아보기 위하여 사용·결합될 수 있는 정보(알고리즘, 매핑테이블 정보, 가명처리에 사용된 개인정보 등)<br>※ 가명처리 과정에서 생성·사용된 정보에 한정된다는 점에서 다른 정보와 구분됨 |
| 재식별 | 특정 개인을 알아볼 수 없도록 처리한 가명정보에서 특정 개인을 알아보는 것 |
| 적정성 검토 | 본 가이드라인에서 제시하고 있는 절차를 기반으로 가명처리 방법 및 수준의 적정성, 가명처리의 적정성, 처리 목적 달성 가능성 등을 검토하는 절차 |
| 가명정보처리시스템 | 개인정보를 가명처리하거나 가명정보를 처리할 수 있도록 체계적으로 구성한 시스템 |

| 구분 | 용어설명 |
| --- | --- |
| 결합키 | 결합 대상 가명정보의 일부로서 해당 정보만으로는 특정 개인을 알아볼 수 없으나 다른 결합대상정보와 구별할 수 있도록 조치한 정보로서, 서로 다른 가명정보를 결합할 때 매개체로 이용되는 값 |
| 결합키연계정보 | 결합키가 동일한 정보에 관한 가명정보를 결합할 수 있도록 서로 다른 결합신청자의 결합키를 연계한 정보 |
| 결합신청자 | 가명정보의 결합을 신청하는 개인정보처리자(공공기관, 법인, 단체, 개인 등)<br>* 가명정보를 제공만 하는 자, 가명정보를 제공하고 결합정보를 이용하는 자, 가명정보를 제공하지는 않지만 결합정보를 이용하는 자를 모두 포함 |
| 결합전문기관 | 보호법 제28조의3 제1항에 따라 서로 다른 개인정보처리자 간의 가명정보 결합을 수행하기 위해 개인정보위 또는 관계 중앙행정기관의 장이 지정하는 전문기관 |
| 결합키관리기관 | 보호법 시행령 제29조의3 제2항에 따라 결합키연계정보를 생성하여 결합전문기관에 제공하는 등 가명정보의 안전한 결합을 지원하는 업무를 하는 한국인터넷진흥원 또는 개인정보위가 지정하여 고시하는 기관 |

# Ⅱ. 가명처리

## 1 개요

- 개인정보처리자(공공기관, 법인, 단체, 개인 등)는 통계작성, 과학적 연구, 공익적 기록보존 등을 위하여 정보주체의 동의 없이 가명정보를 이용, 제공, 결합 등 처리 할 수 있음(보호법 제28조의2 제1항, 제28조의3 제1항)

  ※ (주의) 「가명정보 처리에 관한 특례」에 따라 정보주체의 동의 없이 처리가 가능한 가명정보는 통계작성, 과학적 연구, 공익적 기록보존에 한정되므로 처리 목적이 설정되지 않은 상황에서 보유하고 있는 개인정보를 가명처리하여 보관하는 것은 「가명정보 처리에 관한 특례」에 근거한 처리로 볼 수 없음

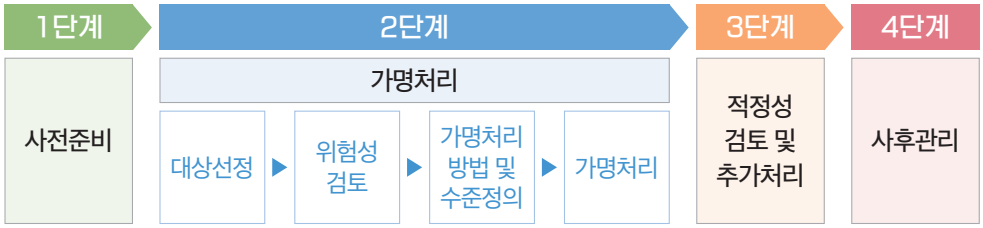
> **제28조의2(가명정보의 처리 등)** ① 개인정보처리자는 통계작성, 과학적 연구, 공익적 기록보존 등을 위하여 정보주체의 동의 없이 가명정보를 처리할 수 있다.
>
> **제28조의3(가명정보의 결합 제한)** ① 제28조의2에도 불구하고 통계작성, 과학적 연구, 공익적 기록보존 등을 위한 서로 다른 개인정보처리자 간의 가명정보의 결합은 보호위원회 또는 관계 중앙행정기관의 장이 지정하는 전문기관이 수행한다.

- 가명정보는 당초 가명처리를 수행한 당시의 처리 목적과 처리 상황(활용 형태, 처리 장소, 방법)에 따라 이용하는 것이 원칙임
  - 다만, 보호법 제28조의2 제1항 및 제28조의3 제1항의 목적 내로 사용하는 경우 가명정보를 당초 처리 목적과 다른 목적으로 이용하거나 제3자로부터 제공받은 가명정보를 다른 제3자에게 재제공하는 등에 제한은 없음

    ※ (예외) 제공 계약 시 재제공 제한이 있거나 반출 시 이용 범위의 제한이 있는 경우에는 가명정보의 재제공 또는 목적 외 이용이 불가할 수 있음

○ 개인정보의 가명처리는 ① 가명처리에 필요한 사전준비, ② 위험성을 검토하여 가명처리 수행, ③ 사전준비와 가명처리 결과에 대한 적정성 검토 및 추가처리, ④ 가명정보를 안전하게 관리하는 사후관리 단계로 이루어 짐

〈 개인정보의 가명처리 단계별 절차도 〉




〈 기타 참고사항 〉

▶ 개인정보처리자는 보다 안전한 가명정보 처리를 위해 다음의 사항을 참고하여 업무에 반영할 수 있음

- 가명처리 관련 업무의 총괄·관리 및 의사결정을 위한 총괄부서(또는 담당자)를 지정할 수 있으며, 주요 업무는 다음과 같음

1) 가명처리 신청(목적)에 대한 적합성 검토
2) 가명처리
3) 가명처리 적정성 검토
4) 가명정보를 처리하는 자에 대한 관리·감독
5) 가명정보에 대한 안전성 확보조치 수행
6) 그 외 안전하고 효율적인 가명정보 처리를 위해 필요한 사항

※ 1), 3)의 경우 외부전문가를 포함한 심의위원회를 구성·운영할 수 있음

▶ 가명처리 관련 업무 담당자의 분리

- 가명처리를 수행한 자와 가명처리의 적정성을 검토하는 자*, 가명정보를 처리하는 자는 관리적·기술적으로 권한을 분리

* 추가정보의 내용을 알고 있는 자가 가명처리의 적정성을 검토를 수행하거나 가명정보를 처리(활용)하는 경우 특정 개인을 알아볼 우려가 있음

# 2 절차

| 1단계 | 2단계 | 3단계 | 4단계 |
|---|---|---|---|
| 사전준비 | 가명처리 | 적정성 검토 및 추가처리 | 사후관리 |

## 1단계 사전준비: 목적 설정 및 가명처리 대상 정보 확보

가명정보 처리 목적을 명확히 하고 가명처리를 위한 적합성 검토 및 계약서, 개인정보 처리방침(48p), 내부 관리계획(85p) 등 필요한 서류를 작성

- 목적 설정: 개인정보처리자는 가명처리 시 법률에서 허용하는 목적 내에서 가명처리하는 목적을 최대한 명확히 설정하여야 함

### ❖ "통계작성"을 위한 가명정보 처리

- "통계"란 특정 집단이나 대상 등에 관하여 작성하는 수량적인 정보를 의미하고, "통계작성을 위한 가명정보 처리"란 통계를 작성하기 위해 가명정보를 이용, 분석, 제공하는 등 처리하는 것을 말함
- 가명정보의 처리 목적이 통계 작성을 위한 것이면 되며 통계의 목적에는 별도의 제한이 없으므로 시장조사를 위한 통계 등 상업적 성격을 가진 통계를 작성하기 위해 가명정보를 처리하는 것도 가능

**예시**

▶ 지방자치단체가 연령에 따른 편의시설 확대를 위해 편의시설(문화센터, 도서관, 체육시설 등)의 이용 통계(위치, 방문자수, 체류시간, 연령, 성별 등)를 작성하고자 하는 경우

### ❖ "과학적 연구"를 위한 가명정보 처리

- "과학적 연구"란 과학적 방법*을 적용하는 연구로서 자연과학, 사회과학 등 다양한 분야에서 이루어질 수 있고, 기초연구, 응용연구 뿐만 아니라 새로운 기술·제품·서비스 개발 및 실증을 위한 산업적 연구도 해당함

  * 과학적 방법이란 체계적이고 객관적인 방법으로 검증 가능한 질문에 대해 연구하는 것을 말함

- "과학적 연구를 위한 가명정보의 처리"란 과학적 연구를 위해 가명정보를 이용, 분석, 제공하는 등 처리하는 것을 말함

- 또한, 보호법 제28조의2 제1항 또는 제28조의3 제1항의 과학적 연구와 관련하여 연구비에 대한 별도의 제한은 법에 없으므로 공적 자금으로 수행하는 연구뿐만 아니라 민간으로부터 투자를 받아 수행하는 과학적 연구에서도 가명정보 처리 가능

**예시**

▶ 코로나19 위험 경고를 위해 생활패턴과 코로나19 감염률의 상관성에 대한 가설을 세우고, 건강관리용 모바일앱을 통해 수집한 생활습관, 위치정보, 감염증상, 성별, 나이, 감염원 등을 가명처리하고 감염자의 데이터와 비교·분석하여 가설을 검증하려는 경우

### ❖ "공익적 기록보존"을 위한 가명정보 처리

- "공익적 기록보존"이란 공공의 이익을 위하여 지속적으로 열람할 가치가 있는 정보를 기록하여 보존하는 것을 의미하고, "공익적 기록보존을 위한 가명정보 처리"란 공익적 기록보존을 위해 가명정보를 이용, 분석, 제공하는 등 처리하는 것을 말함
- 공공기관이 처리하는 경우에만 공익적 목적이 인정되는 것은 아니며, 민간기업, 단체 등이 일반적인 공익을 위하여 기록을 보존하는 경우에도 공익적 기록보존 목적이 인정됨

**예시**

▶ 연구소가 현대사 연구 과정에서 수집한 정보 중 사료가치가 있는 생존 인물에 관한 정보를 기록·보존하고자 하는 경우

### ❖ 민감정보와 고유식별정보의 처리

- 민감정보(보호법 제23조) 또는 고유식별정보(보호법 제24조)도 가명정보 처리 특례에 따라 가명처리하여 활용하는 것이 가능하지만, 개인정보 보호 원칙(보호법 제3조)을 준수하여 처리 목적에 필요하지 않은 민감정보 또는 고유식별정보는 삭제하여야 함
- 다만, 주민등록번호의 경우 법에 명시적으로 주민등록번호를 가명처리 할 수 있는 근거가 없는 한 가명정보 처리 특례에 따른 가명처리는 허용되지 않음(보호법 제24조의2)

※ 가명처리의 목적이 적법한지에 대한 입증 책임은 개인정보처리자에게 있으므로 개인정보처리자는 향후 처리 목적에 대한 증빙을 위해 연구계획서 등 목적설명서를 작성할 수 있음 (「참고자료」 참고 4. 가명처리 및 결합 목적 증빙 자료 예시)

▶ (적절하지 않은 예시) 신제품 개발을 위한 과학적 연구 수행
※ 목적이 구체적으로 명시되지 않아 적절하지 않음

▶ (적절한 예시) ○○제품의 성능 개선을 위해 개인별 ○○○특성에 대한 설문조사를 토대로 개인별 특성과 성능 요인의 연관성에 대한 과학적 연구

- 가명처리 개인정보파일 대상 선정: 처리목적 달성에 필요한 정보의 종류, 범위를 명확히 하여 가명처리 대상을 선정

  ※ 보호법 제37조에 따라 정보주체가 자신의 개인정보에 대한 가명처리 정지를 요구한 대상을 제외하고 선정해야 함

- 가명처리 여부 검토(개인정보 보유부서 또는 가명정보 활용 관련 전담부서 등): 개인정보의 수집 목적 및 성격, 가명정보 활용 목적 등을 고려하여 가명처리 여부를 결정

  ※ 필요시 심의위원회 심사 또는 외부전문가 평가 등을 통해 결정할 수 있음

- 가명정보 처리 상황 정의: 가명처리는 활용 형태, 처리 장소, 방법 등 처리 상황을 고려하여 수행해야하므로 가명처리 전 해당 사항을 미리 확인

- 가명정보 처리를 위한 안전조치: 개인정보 처리방침 수립·공개(보호법 제30조), 내부 관리계획 수립·시행(개인정보의 안전성 확보조치 기준 제4조, 개인정보의 기술적·관리적 보호조치 기준 제3조) 등 가명정보 처리에 앞서 이행하여야 할 사항 등 준비

- 필요서류 작성: 가명정보의 처리 또는 가명처리를 위탁(보호법 제26조에 따라 수행)하거나 가명정보를 제3자에게 제공하는 경우 필요에 따라 재식별 금지에 관한 사항, 기타 처리에 있어 유의해야 할 사항* 등을 포함한 계약서를 작성할 수 있음

  * (예시) 가명정보의 재제공 금지, 가명정보 재식별 금지, 가명정보의 안전성 확보조치, 가명정보의 처리기록 작성 및 보관, 가명정보의 파기, 재식별 시 책임 및 손해배상 등

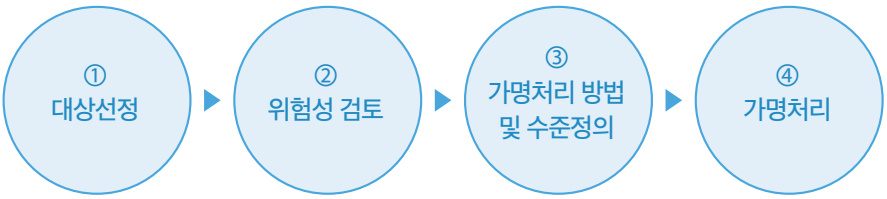
  - 또한, 가명정보 처리 또는 가명처리 위탁 시 위탁 관련 문서 작성, 위탁 업무 공개, 수탁사에 대한 관리·감독에 관한 사항 등 위탁 처리 시 준수하여야 할 사항들을 확인하여야 함

- 기타: 그 밖에 개인정보 활용 및 가명처리 등에 대해 내부 승인 절차를 별도로 두고 있는 개인정보처리자는 이 단계에서 해당 절차를 진행하여야 함

- 가명정보 처리에 관한 내부 관리계획이 없는 경우, 계획 수립 필요
  (「Ⅳ. 가명정보의 안전한 관리」(45p) 참조)

## 2단계 가명처리: 처리 상황에 따른 수준정의 및 처리

가명처리 단계는 세부적으로 ① 대상선정, ② 위험성 검토, ③ 가명처리 방법 및 수준정의, ④ 가명처리를 하는 4가지 단계로 구성되어 있음




- 가명처리 시에는 가명정보 그 자체만으로 특정 개인을 알아볼 수 있는 지와 가명정보를 처리할 자가 보유하고 있거나 접근·입수 가능한 정보*와의 사용·결합을 통해 식별가능한 지를 고려

  * 다른 정보와의 사용·결합을 통해 개인을 식별할 수 있게 되는 경우 보호법 제2조제1호나목에 따른 개인정보에 해당할 수 있음을 주의

**제2조(정의)** 1. "개인정보"란 살아있는 개인에 관한 정보로서 다음 각 목의 어느 하나에 해당하는 정보를 말한다.

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.

**예시**

▶ 전화번호, 지명정보, 소속, 대화 상대방과의 관계 등을 추론할 수 있는 대량의 대화문장 같은 경우 다른 정보와의 사용·결합을 통해 개인을 알아볼 가능성이 있음

- 가명처리 수준은 가명정보 처리 상황에 따라 달라지므로 당초 가명정보를 다른 목적으로 처리하거나 재제공하는 등 가명정보 활용 형태, 처리 장소, 방법 등 처리 상황에 변화가 있는 경우 해당 상황을 고려한 추가적인 가명처리 필요

## 1. 대상선정

- [1단계. 사전준비]에서 설정한 목적을 달성하기 위해 필요한 항목을 개인정보파일에서 선정

  ※ 가명처리 대상항목 선정 시 가명정보 처리 목적 달성에 필요한 최소 항목으로 해야 함

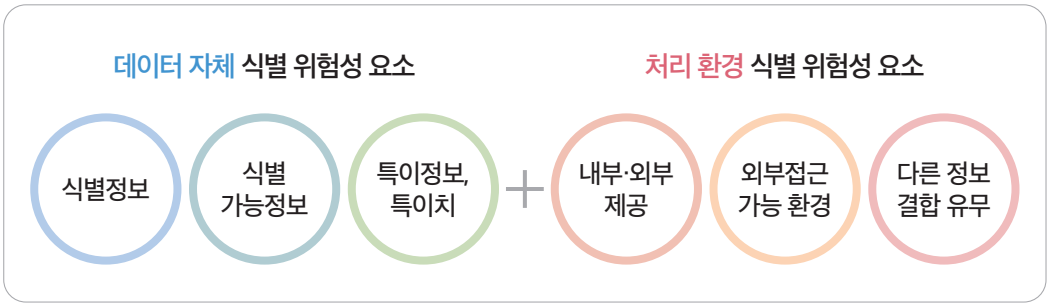
**〈 가명처리 대상 정보 선정(예시) 〉**

- ▶ 개인정보파일 내 항목: 이름, 휴대폰번호, 성별, 이메일, 주소, 구매상품, 구매액, 장바구니 목록
- ▶ 가명처리 목적: 성별과 지역에 따른 구매액 상관관계를 분석하고자 함
- ▶ 가명처리 대상 항목: 성별, 주소(시군구), 구매액

  * 분석 목적과 상관없는 정보는 제외하고 대상 선정

## 2. 위험성 검토

- 위험성 검토는 처리하고자 하는 데이터의 식별 위험성을 가명처리 방법 및 수준에 반영하기 위한 절차이며, 식별 위험성은 1) 데이터 자체의 위험성과 2) 처리 환경의 위험성으로 구분하여 검토




### 1) 데이터 자체의 위험성 검토

- 데이터 자체 위험성 검토는 가명처리의 대상이 되는 정보에 식별 가능한 요소가 있는지를 파악하는 것으로, 그 자체로 식별될 위험이 있는 항목, 다른 항목과 결합을 통해 식별될 가능성이 있는 항목, 그 밖에 특이정보, 특이치 등이 있는지 검토

  - (이용 항목별 식별 위험) 다른 사람과 구분하기 위해 부여된 식별 정보는 특정 개인과 고유하게 연결되어 있으므로, 해당 정보가 포함되어 있을 경우 특정 개인을 알아볼 가능성이 높음

〈 개인 식별 가능성이 높은 정보(예시) 〉

▶ 식별정보
- 고유식별정보(여권번호, 외국인등록번호, 운전면허번호), 성명, 전화번호, 전자우편주소, 의료기록번호, 건강보험번호 등 식별을 목적으로 생성된 정보

▶ 식별가능정보
- 성별, 연령(나이), 거주 지역, 국적, 직업, 위치정보 등 개인정보처리자의 입장에서 개인을 알아볼 수 있는* 정보

* 개인을 '알아볼 수 있는지'는 해당 정보를 처리하는 자(정보의 제공 관계에 있어서는 제공받는 자를 포함)를 기준으로 판단하여야 함

- (다른 이용 항목과의 결합 유무) 단일 이용 항목으로는 식별 가능성이 없으나, 가명처리 대상이 되는 다른 이용 항목과 결합하여 식별 가능성이 높아지는 이용 항목이 있는지 검토

※ 지역-직업-나이, 직장-직위 등

- (특이정보, 특이치 유무) 가명처리 대상 전체 데이터에 식별 가능성을 가지는 고유(희소)한 값이 있는지, 편중된 분포를 가지는 단일·다중 이용 항목이 있는지 검토

〈 특이정보(예시) 〉

▶ 특이정보
- 희귀 성씨, 희귀 혈액형, 희귀 눈동자 색깔, 희귀 병명, 희귀 직업 등 정보 자체로 특이한 값을 가지고 있는 정보
- 국내 최고령, 최장신, 고액체납금액, 고액급여수급자 등 전체적인 패턴에서 벗어나 극단값이 발생할 수 있는 정보

**2) 처리 환경의 위험성 검토**

● 개인정보처리자는 가명정보 활용 형태, 처리 장소, 방법 등 가명정보 처리 상황에 따라 발생할 수 있는 식별 위험성 검토

- (내부이용 및 외부제공 여부) 처리 주체가 보유하고 있는 정보 또는 접근·입수 가능한 정보와 이용 범위 및 유형을 고려하여 식별가능한 항목이 있는지 검토

※ 처리 주체가 보유, 접근, 입수 가능한 모든 정보를 고려하여 식별 가능성을 검토할 필요는 없으며, 보안서약서, 계약서 등을 통해 파악이 가능한 범위의 정보를 고려하여 식별 가능성 검토 가능

- (처리 장소) 가명정보가 해당 가명정보 외에 다른 정보의 접근·입수가 제한된 장소에서 처리되는지 검토

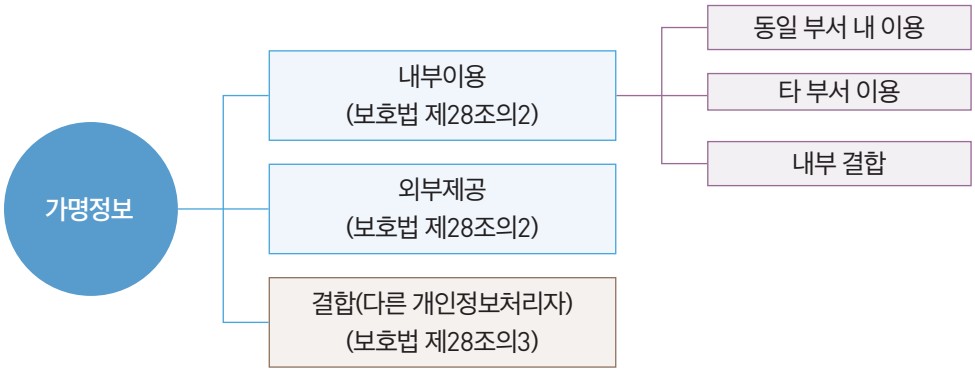
  ※ 다만, 보안서약서, 계약서 등으로 내·외부 정보의 활용이 제한된 경우 폐쇄 환경에 준하여 검토 가능

- (다른 정보 결합 유무) 가명정보를 다른 정보와 연계 분석할 예정인 경우 다른 정보와 결합하여 식별가능한 항목이 있는지 검토

- 가명정보를 다른 정보와 내부 결합 할 예정인 경우 다른 정보와 결합하여 식별가능한 항목이 있는지 검토

- 가명정보를 반복 제공할 예정인 경우 반복 제공을 통해 식별 위험이 높아지는 항목이 있는지 검토

● 가명정보 이용 및 제공시 유의 사항




### 가. 내부이용

▷ 개인정보처리자가 보유(정보주체로부터 직접 수집하거나 합법적으로 수집·제공받은 개인정보)한 개인정보를 가명처리 또는 내부 결합하여 직접 활용 또는 다른 부서에 제공하는 경우를 의미

▷ 가명정보를 처리하는 소속 부서에서 이미 보유하고 있는(접근 가능한) 정보 및 처리 시점을 기준으로 제공받는 다른 정보를 고려하여 식별 위험성을 검토하여야 함

### 잘못된 내부이용(동일 부서 내 이용) 사례

동일 부서 내 이용으로 ○○화장품 회사의 A팀은 화장품 판매정보를 관리하는 팀으로서, 가명정보 또는 추가정보에 접근할 수 있는 권한을 분리하지 않고 해당 정보를 가명처리하여 신상품 수요조사 예측 모델 개발을 목적으로 활용

☑ (처리현황) A팀은 판매정보 내 개인식별 가능성이 있는 이름, 성별, 승인번호를 가명처리하고, 희귀 지역의 판매내역을 삭제하여 A팀 가명정보 분석담당자에게 제공

→ 가명정보 분석담당자는 A팀의 판매정보 관리 업무를 병행하여 업무를 수행하고 있음

☑ (문제점) 가명정보 분석담당자는 가명정보 분석을 통해 최고가 화장품의 금액과 판매지역을 파악할 수 있으며, 판매정보 관리 업무를 병행하고 있어 해당 금액과 지역을 통해 특정 개인을 식별할 가능성이 있음

☑ (해결방안) 가명정보 분석담당자가 가명정보 분석을 수행하는 경우를 제외하고는 특정 개인을 알아볼 수 있는 개인정보처리시스템에 접근할 수 없도록 제한해야 함

### 잘못된 내부이용(타 부서 내 이용) 사례

타 부서 이용으로 □□공사는 A부서의 고속도로 이용차량 빅데이터 분석 결과를 고속도로 통행요금을 관리하는 B부서에 교통서비스 개선을 위한 연구 목적으로 제공(이 때 B부서에서 처리하는 개인정보를 고려하지 않음)

☑ (처리현황) A부서는 개인식별 가능성이 있는 차량번호, 차종 등을 가명처리하고, 이동시간, 이동량, 사고정보 등의 정보를 B부서에 제공

→ B부서는 고속도로 통행요금 관리를 위해 고객번호와 차량번호, 톨게이트 입출시간 및 결제금액 정보를 보유하고 있음

☑ (문제점) B부서는 A부서에서 제공받은 정보의 이동시간 정보와 B부서가 보유한 톨게이트 입출시간을 활용하여 특정시간에 통과한 차량의 번호를 알 수 있으며, 해당 차량번호를 통해 특정 개인을 식별할 가능성이 있음

☑ (해결방안) A부서에서는 B부서가 보유하고 있는 정보를 고려하여 특정 시간에 대한 식별 가능성이 없도록 이동시간 삭제 또는 가명처리 등을 수행하여야 함(필요 시 가명처리를 위해 B부서가 보유한 톨게이트 입출시간 정보 제공 요청)

## 나. 외부제공

▷ 개인정보처리자가 보유한 개인정보를 가명처리하여 특정 제3자에게 제공하는 경우를 의미

- 제3자의 개인정보 보호수준 및 신뢰도를 고려하여 가명정보 제공으로 인하여 발생할 수 있는 재식별 위험을 최소화하기 위하여 노력하여야 함*

  * 보호수준이 낮은 기관에는 상대적으로 높은 수준의 가명처리 수준을 적용하는 방법 등

- 가명정보를 제3자에게 제공하는 경우 추가정보 등 특정 개인을 알아보기 위하여 사용될 수 있는 정보를 제공하여서는 아니됨(보호법 제28조의2 제2항)

> **제28조의2(가명정보의 처리 등)** ② 개인정보처리자는 제1항에 따라 가명정보를 제3자에게 제공하는 경우에는 특정 개인을 알아보기 위하여 사용될 수 있는 정보를 포함해서는 아니 된다.

- 또한, 개인정보처리자는 제3자가 사전에 보유하고 있는 정보 및 처리 시점을 기준으로 제공받는 다른(개인)정보 등을 고려하여야 하고, 이를 파악하기 위해 관련 정보*를 요청하는 것도 가능

  * 제3자가 관리하고 있는 개인정보 중 제공받는 가명정보와 결합 가능성이 있는 개인정보 목록 등

- 사전준비 단계의 계약서에 재식별에 대하여 명시한 사항이 있다면 이를 고려할 수 있음

**잘못된 외부제공 사례**

○○호텔에서는 최고급 객실을 이용한 VIP등의 특이정보를 삭제하지 않고 호텔 투숙 및 서비스 금액 등을 △△분석 회사에 제공하고, △△분석회사는 해당 정보를 분석하여 시간에 따른 객실 이용현황 및 서비스이용에 대한 조사 연구를 수행

☑ (처리현황) △△분석회사는 온라인 SNS정보 및 다양한 기업의 정보를 수집하여 다양한 연구조사를 실시하는 회사로서 내부 관리계획을 수립하고, 관리적·기술적 보호조치를 준수하고 있음

→ ○○호텔은 회원번호와 이름을 가명처리하고, 나이, 성별, 등급, 예약방법, 객실정보, 체크인, 체크아웃, 서비스 이용금액을 제공

☑ (문제점) △△분석회사의 분석담당자는 특정일에 최고급 객실을 이용한 내용을 분석과정에서 인지할 수 있으며, 기존 업무(온라인 SNS정보 수집)를 수행하며 공개된 정보(예: 개인이 SNS에 올리는 정보, 여행후기 등)를 통해 특정 개인을 식별할 가능성이 있음

☑ (해결방안) ○○호텔은 제공하는 가명정보에 포함된 특이정보(최고급 객실)를 삭제 또는 가명처리 등을 수행하여야 함

○ 개인정보처리자는 데이터 자체의 위험성과 처리 환경의 위험성 검토를 통해 가명처리에 대한 식별 위험성 평가 결과를 도출하여야 함

〈 식별 위험성 검토 체크리스트 예시 〉

| 구분 | | 식별 위험성 검토 항목 검토 |
|---|---|---|
| 데이터 | 식별성 | 개인 식별이 가능한 항목이 포함되어 있는가 |
| | | 예시<br>▶ 식별을 목적으로 한 단일항목의 정보가 있는가<br>▶ 두 개 이상의 컬럼(항목) 조합하여 식별 가능성이 높아지는 정보가 있는가<br>▶ 공개된 데이터와 결합·대조하여 식별 가능성이 높아질 수 있는 이용 항목이 있는가<br>- 예시) 통계청의 인구 센서스 데이터를 사용하여 식별가능한 이용 항목이 있는가 |
| | 특이정보 및 특이치 | 데이터 분포가 편중되어 있어 식별 가능성이 있는 이용 항목이 있는가 |
| | | 예시<br>▶ 연속적인 숫자형 데이터에서 데이터 값의 분포가 양 끝단의 정보(분포 곡선에 따라 한쪽의 정보 포함)가 현저히 낮은 항목이 있는가<br>▶ 일반적인 문자형 데이터(비 연속적인 숫자형 데이터 및 코드형 데이터 포함)에서 특정 값으로 현저히 낮은 항목이 있는가 |
| 처리환경 | 이용 및 제공 | 처리주체가 보유하고 있는 정보 또는 접근·입수 가능한 정보와 이용 범위 및 유형을 고려하여 식별가능한 항목이 있는가 |
| | | 예시<br>▶ 시계열 분석 등을 위한 목적으로 가명정보를 반복 제공할 예정인 경우 반복 제공을 통해 식별 위험이 높아지는 항목이 있는가<br>▶ 가명정보 제공 시 개인정보처리자의 개인정보 보호 수준 및 신뢰할 수 있는 수준을 고려하였는가 |
| | 처리장소 | 가명정보의 처리가 다른 정보를 접근·입수할 수 있는 장소에서 처리 가능한가 |
| | | 예시<br>▶ 외부접근이 제한되어 있는 폐쇄망 형태의 장소인가 |
| | 다른 정보 결합 | 처리주체가 보유하거나 접근·입수 가능한 정보 등 다른 정보와 연계 또는 결합하여 식별가능한 항목이 있는가 |

〈 식별 위험성 검토 결과보고서 예시 〉

| | | |
|---|---|---|
| 가명정보 활용목적 | | ▶ A사가 보유한 부동산 시세정보를 가명처리하여 B기관에 제공하여, 부동산 임대소득 계산 및 인근지역 시세자료 파악을 위한 연구 수행 |
| 가명처리 대상 데이터 항목 | | ▶ 소유자명, 연락처, 주택구분, 시도, 시군구, 읍면동, 지번, 전용면적, 공급면적, 전세, 보증금, 월세<br>※ 항목을 나열하지 못하는 경우 '별지' 사용 가능 |
| 처리 환경 검토 | 가명정보 이용 및 제공 형태 | ▶ 특정 제3자(B기관) 제공<br>- A사는 B기관과 계약체결을 통해 가명정보를 제공 |
| | 제공받는 자의 처리 환경 | ▶ 가명정보를 제공받는 B기관은 부동산 관련 다른(개인)정보를 보유하고 있지 않으며, 다른 정보에 접근 및 입수할 수 없도록 접근이 제한된 장소에서 수행 예정 |
| | 제공 받는 자의 개인정보 보호 수준 | ▶ B기관은 개인정보(가명정보)처리시스템에 대한 ISMS-P인증을 취득하고 있으며, 내부 관리계획 수립 및 시행을 통해 개인정보(가명정보)의 안전한 관리를 수행하고 있음 |
| 데이터 항목별 위험성 분석 | | ▶ '소유자명', '연락처'는 식별정보,<br>▶ '지번'은 식별가능정보, '시세정보(전세, 보증금, 월세)'는 특이정보 가능성 존재 |
| 최종 검토의견* | | ▶ 해당 연구는 특정 제3자와의 계약 체결을 통해 가명정보를 활용하는 경우에 해당하며, 제공받는 자가 별도의 다른(개인)정보를 통해 가명정보를 재식별 할 가능성이 낮음<br>- '소유자명', '연락처'는 활용 목적상 반드시 필요한 경우가 아니라면 삭제 또는 목적 달성에 필요한 경우 가명처리 필요<br>- '지번' 및 '시세정보(전세, 보증금, 월세)'의 경우 다른(공개된 정보 등) 정보를 통해 재식별 가능성이 있어 삭제 또는 목적 달성에 필요한 경우 가명처리 필요<br>▶ 그 외의 정보들은 재식별 가능성이 낮으며 목적 달성을 위해 필요하다고 판단되므로 가명처리하지 않음 |

※ 최종 검토의견은 외부전문가에게 자문 및 작성을 요청할 수 있음

## 3. 가명처리 방법 및 수준정의

- 개인정보처리자는 '식별 위험성 검토 결과보고서'를 기반으로 가명정보의 활용 목적 달성에 필요한 수준을 고려하여 가명처리 방법 및 수준 정의를 하여야 함

  ※ 가명처리 기법 등은 (참고자료 개인정보 가명처리 기술 및 예시 54p 참조)

〈 **가명처리 방법 및 수준 정의표 예시** 〉

○ '식별 위험성 검토 결과보고서'에서 분류한 개인정보에 대한 가명처리 수준 정의

| 순번 | 항목명 | 처리수준 | 비고 |
|---|---|---|---|
| 1 | 소유자명 | - 가명처리<br>(암호화: SHA2+Salt) | - 소유자명과 연락처는 추후 시계열 분석을 위해 가명처리 수행 |
| 2 | 연락처 | | |
| 3 | 지번 | - 가명처리(삭제) | - 세부 지번의 정보는 분석목적에 필요하지 않음 |
| 4 | 전세 | - 기타기술<br>(라운딩: 만원 단위) | - 만원 단위의 금액만 분석목적에 필요 |
| 5 | 보증금 | | |
| 6 | 월세 | | |
| 7 | 주택구분 | - 처리하지 않음<br>※ 항목이 다수여서 작성이 어려운 경우 '별지'를 활용하여 목록만 제시 | - 처리하지 않는 항목을 작성 |
| 8 | 시도 | | |
| 9 | 시군구 | | |
| 10 | 읍면동 | | |
| 11 | 전용면적 | | |
| 12 | 공급면적 | | |

## 4. 가명처리

- 개인정보처리자는 '가명처리 방법 및 수준 정의표'를 기반으로 가명처리를 수행하여야 함

〈 가명처리 절차 (예시) 〉

### (원본정보)

| 소유자 명 | 연락처 | 주택 구분 | 법정동코드 | 시도 | 시군구 | 읍면동 | 지번 | 건물명 | 전세 (천원) | 보증금 (천원) | 월세 (천원) | 전용 면적 | 공급 면적 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 김철수 | 090-1234-5678 | 아파트 | 2635010700 | 서울 특별시 | 동작구 | 사당동 | 1388-4 | 대우 마리나 | - | 25,000 | 750 | 104.00 | 84.00 |
| 이영희 | 090-2468-3579 | 오피 스텔 | 3611011000 | 대전 광역시 | 서구 | 둔산동 | 656 | 푸르 지오 시티 | 81,250 | - | - | 56.45 | 24.32 |
| 박민호 | 090-9876-5432 | 아파트 | 4311410100 | 부산 광역시 | 해운대구 | 우동 | 111-13 | 평화 | 125,000 | - | - | 100.00 | 84.00 |

선정 선정 선정

### (대상선정)

- 목적 : 부동산 임대소득 계산 및 인근지역 시세자료 파악을 위한 연구

| 소유자 명 | 연락처 | 주택 구분 | 시도 | 시군구 | 읍면동 | 지번 | 전세 (천원) | 보증금 (천원) | 월세 (천원) | 전용 면적 | 공급 면적 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 김철수 | 090-1234-5678 | 아파트 | 서울 특별시 | 동작구 | 사당동 | 1388-4 | - | 25,000 | 750 | 104.00 | 84.00 |
| 이영희 | 090-2468-3579 | 오피 스텔 | 대전 광역시 | 서구 | 둔산동 | 656 | 81,250 | - | - | 56.45 | 24.32 |
| 박민호 | 090-9876-5432 | 아파트 | 부산 광역시 | 해운대구 | 우동 | 111-13 | 125,000 | - | - | 100.00 | 84.00 |

### (위험성 검토)

- 데이터 위험성과 처리 환경의 위험성 검토에 따라 가명처리 방법 및 수준 정의

**데이터**

소유자명, 연락처는 개인정보로 분류하고 가명처리(암호화)

구체적인 지번은 분석목적에 관계 없어 삭제조치 및 시세정보는 분석에 필요한 단위(만원)로 가명처리

+

**처리환경**

A사의 부동산 시세정보를 B기관에 제공(계약)

제공되는 항목의 '지번'의 경우 등기부열람을 통해 특정개인식별 가능성 존재

식별정보 / 식별가능정보

| 소유자 명 | 연락처 | 주택 구분 | 시도 | 시군구 | 읍면동 | 지번 | 전세 (천원) | 보증금 (천원) | 월세 (천원) | 전용 면적 | 공급 면적 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 김철수 | 090-1234-5678 | 아파트 | 서울 특별시 | 동작구 | 사당동 | 1388-4 | - | 25,000 | 750 | 104.00 | 84.00 |
| 이영희 | 090-2468-3579 | 오피스텔 | 대전 광역시 | 서구 | 둔산동 | 656 | 81,250 | - | - | 56.45 | 24.32 |
| 박민호 | 090-9876-5432 | 아파트 | 부산 광역시 | 해운대구 | 우동 | 111-13 | 125,000 | - | - | 100.00 | 84.00 |

(소유자 명, 연락처)
+Salt
암호화

삭제 / 라운딩

(가명처리)

| ID | 주택 구분 | 시도 | 시군구 | 읍면동 | 전세 (천원) | 보증금 (천원) | 월세 (천원) | 전용 면적 | 공급 면적 |
|---|---|---|---|---|---|---|---|---|---|
| wd4e85D2C1qe89rwqe | 아파트 | 서울 특별시 | 동작구 | 사당동 | - | 25,000 | 800 | 104.00 | 84.00 |
| r5w1e2SXzi4wd64qwz | 오피스텔 | 대전 광역시 | 서구 | 둔산동 | 81,300 | - | - | 56.45 | 24.32 |
| ghe6W15Z5ax4Qe24jx | 아파트 | 부산 광역시 | 해운대구 | 우동 | 125,000 | - | - | 100.00 | 84.00 |

- 가명처리 단계에서 생성되는 추가정보는 원칙적으로 파기하고 필요한 경우 가명정보와 분리하여 별도로 저장하여야 함

  - 추가정보의 분리보관은 [IV.가명정보의 안전한 관리] 기술적 보호조치 49p 참조)

## 3단계 적정성 검토 및 추가 가명처리

- 가명처리가 적절한 수준으로 이루어 졌는지에 대한 최종적인 판단절차를 수행하여야 함
  - [2단계. 가명처리]에 따라 가명처리가 되었는지 확인하고, 가명처리 한 결과가 가명정보의 처리 목적을 달성하기 위해 적합한지 검토
  - 만약 재식별 가능성이 있다고 판단한 경우 [2단계. 가명처리]를 반복하거나 부분적으로 추가적인 가명처리를 수행

    ※ 데이터의 분포, 내용 등을 검토하여 특이정보가 추가로 생성 또는 발견된 경우 재식별 가능성을 낮추기 위한 적절한 조치를 취하여야 함

- 가명처리에 대한 적정성 검토는 개인정보처리자의 판단에 따라 내부 인원을 활용하여 자체적으로 검토할 수 있으며, 필요시 외부전문가를 통하여 검토할 수 있음

### 1. 적정성 검토

- (필요서류 및 위험성 검토) 필요서류 내용, 데이터 자체 위험도, 처리환경 등 위험성 판단 항목을 누락 없이 검토하였는지 확인

  ※ 사전준비 단계에서 필요서류가 법/제도 목적에 적법하게 작성되었는지와 가명처리 단계에서 체크리스트 및 결과보고서 기반으로 위험성 판단 항목을 누락 없이 검토하였는지 여부

- (가명처리 방법 및 수준의 적정성) 가명처리 단계에서 위험성 검토 결과를 반영하여 가명처리 방법 및 수준을 적정하게 정의하였는지 확인

- (가명처리의 적정성) 정의한 가명처리 방법 및 수준에 따라 실제 가명처리를 수행하였는지 확인

  ※ 특히 대용량 정보의 경우 중간에 처리되지 않은 부분이 있을 수 있으므로 가능한 가명정보 항목 전체를 확인 필요

- (처리 목적 달성 가능성) 가명처리를 수행한 정보가 당초 가명정보 처리 목적을 달성할 수 있는지 여부 검토

  ※ 목적 달성에 필요한 최소한의 항목으로 처리되었는지와 처리된 정보가 당초 목적을 달성하기에 적정한지 판단

### 2. 추가 가명처리

- 적정성 검토 결과 가명처리가 적정하지 않다고 판단되면 가명처리를 다시 수행하거나 부분적으로 추가적인 가명처리를 수행할 수 있음

## 4단계 사후관리

- 적정성 검토 결과 가명처리가 적정하다고 판단되면 가명정보를 본래 처리 목적을 위해 활용할 수 있으며, 법에 따라 기술적·관리적·물리적 안전조치 등 사후관리를 이행하여야 함

  ※ 구체적 내용은 「Ⅳ. 가명정보의 안전한 관리」(45p) 참조 참고

### 1. 재식별 금지 및 모니터링

- 개인정보처리자는 누구든지 특정 개인을 알아보기 위한 목적으로 가명정보를 처리해서는 아니 되며(보호법 제28조의5 제1항), 가명정보를 처리하는 중 우연히 특정 개인이 식별되는 경우 즉시 처리중지, 회수, 파기 등 위와 같은 위험을 제거하기 위해 적절한 조치를 수행하여야 함(보호법 제28조의5 제2항)

> **제28조의5(가명정보 처리 시 금지의무 등)** ① 누구든지 특정 개인을 알아보기 위한 목적으로 가명정보를 처리해서는 아니 된다.
> ② 개인정보처리자는 가명정보를 처리하는 과정에서 특정 개인을 알아볼 수 있는 정보가 생성된 경우에는 즉시 해당 정보의 처리를 중지하고, 지체 없이 회수·파기하여야 한다.

- 또한, 개인정보처리자는 가명정보 처리 과정에서 특정 개인이 식별될 위험이 있는지 여부를 지속적으로 모니터링 하여 안전하게 처리하여야 함

### 2. 안전조치 시행

- 개인정보처리자는 사전준비 단계에서 수립한 내부 관리계획에 따라 가명정보를 안전하게 관리하여야 함

### 3. 개인정보 처리방침 수립 및 공개

- 개인정보처리자는 가명정보 처리와 관련하여 처리 목적, 처리하는 개인정보의 항목 등을 개인정보 처리방침에 공개하여야 함

### 4. 가명정보 처리 관련 기록 작성 및 보관

- 개인정보처리자는 가명정보의 처리 목적, 개인정보 항목, 이용내역, 제3자 제공 시 제공받는 자를 작성하여 보관하여야 함

〈 기타 참고사항: 내부 결합 〉

● 개인정보처리자는 자신이 보유하고 있는 가명정보를 결합하여 활용할 수 있으며, 결합 절차가 정해져있지는 않지만 결합 과정에서 특정 개인을 알아볼 수 없도록 유의하여 결합을 수행하여야 함

※ 안전한 결합을 위해 결합키를 이용한 결합방법을 선택할 수 있음

〈 가명정보 내부 결합 절차도 〉


- 개인정보처리자는 결합된 정보를 활용할 때 특별한 사유(시계열 분석 등)가 없는 한 결합키 등 결합을 위해 사용한 정보를 삭제한 후 활용하여야 함

※ (주의) 결합키 생성에 이용된 알고리즘, 매핑테이블 등은 추가정보에 해당하므로, 결합된 가명정보와 분리하여 보관하여야 하고, 접근권한을 분리하여야 함

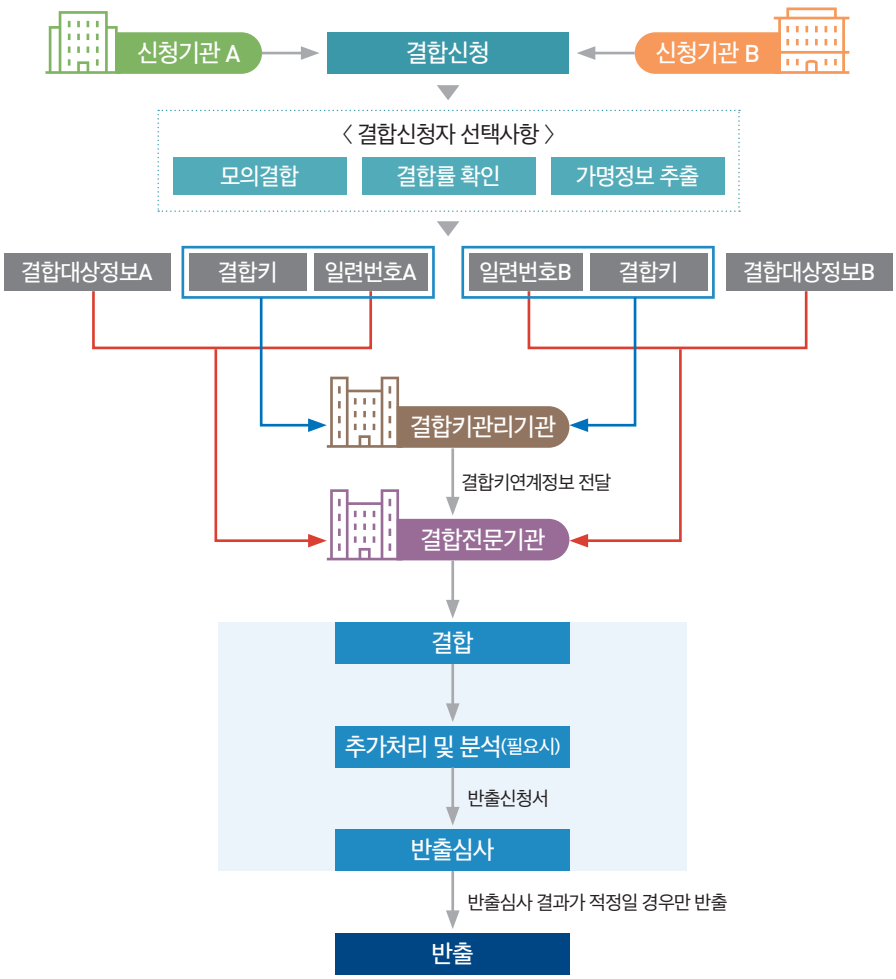
〈 가명정보 내부 결합 방식 예시 〉


※ 결합키 생성 등의 구체적인 내용은 「가명정보 결합 및 반출 절차 」35p 참고

# Ⅲ. 가명정보 결합 및 반출

## 1 개요

〈 가명정보 결합·반출 절차 〉



## 1. 가명정보의 결합

- 개인정보처리자는 결합전문기관을 통해 통계작성, 과학적 연구, 공익적 기록보존 등을 위한 가명정보 결합 가능
  - 서로 다른 개인정보처리자가 보유한 가명정보를 결합하여 활용하고자 하는 경우에는 개인정보위 또는 관계 중앙행정기관의 장이 지정한 결합전문기관을 통하여 수행하여야 함 (보호법 제28조의3 제1항)

## 2. 가명정보의 결합 방식

- 가명정보의 결합은 보통 1회로 종료되지만, 처리목적을 달성하기 위해 필요한 경우 동일한 서로 다른 개인정보처리자 간의 가명정보를 반복적으로 결합*할 수도 있음

  * 시계열 분석 등 처리목적 달성을 위해 정보의 지속적·주기적 결합·분석이 필요한 경우

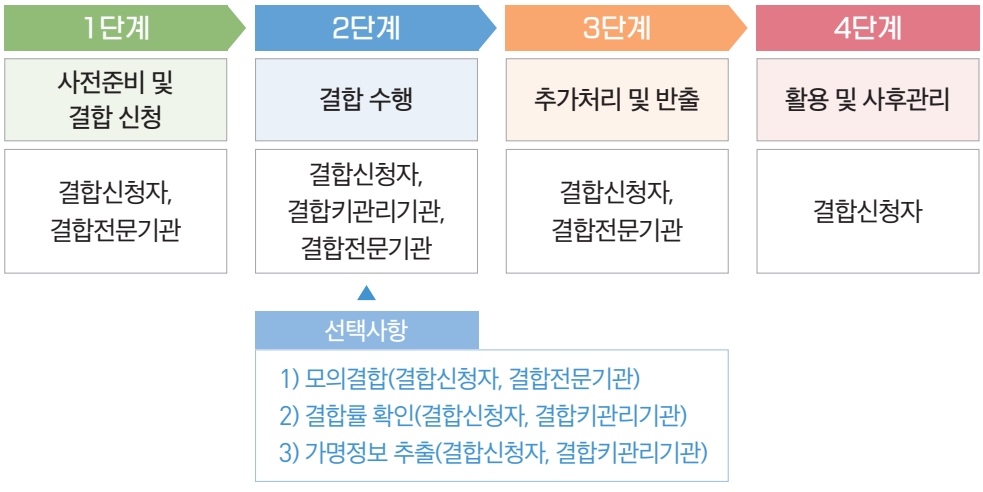
  - 반복결합을 신청한 경우 결합절차에 차이는 없으나, 반복적인 분석을 위해 반출정보에 필요한 키(반복결합키)가 추가로 포함됨

  ※ 시계열 분석을 위한 반복결합 절차의 구체적인 내용은 75p 참고

## 3. 가명정보 결합·반출 절차

- 가명정보 결합·반출은 ① 사전준비 및 결합 신청, ② 결합 수행, ③ 추가처리 및 반출, ④ 활용 및 사후관리의 총 4단계를 거쳐 진행
  - 결합신청자*는 필요에 따라 모의결합, 결합률 확인, 가명정보 추출을 신청하여 수행할 수 있음

    * 가명정보를 보유하고 있는 개인정보처리자, 현재 가명정보를 보유하고 있지 않으나 결합된 가명정보를 처리할 예정인 자



- 1단계: 결합신청자 간의 결합신청에 필요한 사항*을 협의하여 결합신청서를 작성하는 등 가명정보 결합에 필요한 사전 준비사항을 수행하고 그 결과를 반영하여 결합전문기관에 결합 신청

  * 개인정보파일에 가명정보 결합 목적 달성에 필요한 항목 선정, 시계열 분석 여부, 모의결합/결합률 확인/가명정보 추출 신청여부, 결합키 생성항목 등

  - 결합신청자는 결합전문기관과 결합일정, 전송방법 등을 협의

- 2단계: 가명정보를 제공하고자 하는 결합신청자는 결합키관리기관으로부터 결합키 생성에 이용되는 정보(Salt값)를 수신하여 결합키를 생성하고 필요시 모의결합, 결합률 확인, 가명정보 추출 등을 수행한 후 결합에 필요한 정보를 각 기관에 전송

  - 결합신청자는 결합 수행을 위해 아래와 같이 각 기관에 결합에 필요한 정보 전송
    - 결합키관리기관: [결합키+일련번호]
    - 결합전문기관: [결합대상정보+일련번호]

- 3단계: 결결합정보를 이용하고자 하는 결합신청자는 결합정보를 반출하기 전 결합전문기관 내에 설치된 별도의 공간에서 추가 가명·익명처리를 하거나, 결합전문기관이 분석기능을 지원하는 경우 분석을 수행할 수 있으며, 결합정보 또는 분석결과 등을 반출하고자 하는 경우 결합전문기관에 반출신청서를 제출하여 반출 신청

- 4단계: 결합정보를 이용하고자 하는 결합신청자는 반출정보를 당초 결합신청서 및 반출신청서에 기재한 목적에 따라 처리할 수 있으며, 가명정보 처리 시 안전조치를 준수

# 2 절차

○ 가명정보 결합 시 결합 수행에 대한 수행주체(결합신청자, 결합키관리기관, 결합전문기관)별 세부 절차는 다음과 같으며, 본 가이드라인에서는 결합신청자 기준으로 결합절차를 안내함

| 절차 | | 결합신청자 | 결합키관리기관 | 결합전문기관 |
|---|---|---|---|---|
| 1 사전준비 및 결합 신청 | | ① 결합 신청 | - | ② 결합신청서 검토 및 접수<br>③ 결합 일정·절차 등 협의(결합신청자) |
| 2 결합 수행 | 1. 결합키 생성 | ① 결합키 생성 협의<br>③ 결합키 및 일련번호 생성 | ② 결합키 생성 협의(Salt값 전송) | - |
| | 2. 모의결합 (선택) | ① 결합키 전송 | - | ② 모의결합 가능성 검토 및 통지<br>③ 모의결합 대상 결합키 선정 및 전송 |
| | | ④ 모의결합 대상 가명처리<br>⑤ 가명처리된 모의결합대상정보, 가명처리내역 및 결합키 전송 | - | ⑥ 가명처리 수준 검토 (필요시 추가처리 요청)<br>⑦ 모의결합 수행 |
| | | ⑧ 모의결합된 정보 분석 (결합전문기관 내)<br>* 반출 제한 | - | ⑨ 모의결합 관련 정보 파기 |
| | 3. 결합률 확인(선택) | ① 결합키 및 일련번호 전송<br>④ 결합률 확인 | ② 결합키연계정보 생성<br>③ 결합률 측정 및 통보 | - |
| | 4. 가명정보 추출(선택) | ① 결합키 및 일련번호 전송<br>② 추출 요청 | ③ 추출 가능 여부 검토 및 통지<br>④ 추출에 필요한 일련번호 선정 및 전송 | - |
| | 5. 가명처리 및 검토 | ① 결합대상정보 확정<br>② 가명처리<br>③ 가명처리된 결합대상정보, 가명처리 내역 및 일련번호 전송 | - | ★ 가명처리 지원(가능한 경우)<br>④ 가명처리 수준 검토 (필요시 추가처리 요청) |
| | 6. 결합 | ① 결합키 및 일련번호 전송 | ② 결합키연계정보 생성 및 전송<br>* 반복결합의 경우 반복결합키 포함 | ③ 결합키연계정보 수신 및 가명정보 결합<br>* 반복결합의 경우 반복결합키 포함 |
| 3 추가처리 및 반출 | 1. 추가처리 (필요시) | ① 결합된 정보의 추가처리 및 분석(결합전문기관 내)<br>* 결합전문기관에 지원 요청 가능 | - | ★ 추가처리 및 분석 지원 (가능한 경우) |
| | 2. 반출 | ① 반출신청 | ④ 결합키 및 결합키연계정보 파기(필요시 유지*)<br>* 반복결합의 경우, 결합키 생성방법(Salt)·반복결합키 생성방법(Salt) 보관 | ② 반출심사위원회 구성·운영<br>③ 반출 승인 및 결합정보 반출<br>* 반복결합의 경우 반복결합키 포함<br>④ 결합키연계정보 파기 |
| 4 활용 및 사후관리 | | ✓ 안전성 확보 조치 이행<br>✓ 가명정보 처리 내역 기록·보관 | - | ★ 반출한 정보 분석 지원 (가능한 경우)<br>★ 개인정보 보호 교육 제공 (가능한 경우) |

★: 보호법 제11조의2에 따른 결합전문기관의 업무지원 사항으로, 결합신청자는 결합전문기관이 해당 업무에 대해 지원 가능한 경우 요청할 수 있음

## 1단계 사전준비 및 결합 신청

### 1. 사전준비

- 결합신청자는 서로 다른 결합신청자 간의 협의를 통해 가명정보 결합에 대한 사전준비를 수행하여야 함

  ※ 협의사항: 개인정보파일에서 가명정보 결합 목적 달성에 필요한 항목 선정, 시계열 분석 여부, 모의결합/결합률 확인/가명정보 추출 신청여부, 결합키 생성항목 등

  - 결합신청자들은 공통으로 보유하고 있는 정보 중에서 결합키를 생성할 때 활용할 항목을 결정하여야 함

> ▶ 결합키 생성 항목 정의(예시)
> - A사 : 성명, 전화번호, 생년월일, 주소, 차량 정보, 배기량, 주유금액 등
> - B사 : 성명, 전화번호, 생년월일, 주소, 주거형태, 보증금 유무, 월세 유무 등
> ⇨ A, B사가 동일하게 가지고 있는 성명, 전화번호, 생년월일을 결합키 생성 항목으로 선정

- 결합신청자는 필요시 가명정보 결합에 관한 별도의 내부승인절차 등을 이행할 수 있으며, 결합 건에 대한 계약 체결 등 필요한 조치를 할 수 있음

### 2. 결합 신청

- (결합 신청) 결합신청자는 가명정보를 보유하고 있는 개인정보처리자 뿐만 아니라, 현재 가명정보를 보유하고 있지 않으나 결합된 가명정보를 처리할 예정인 개인정보처리자도 결합신청자가 됨

  ※ 결합전문기관은 결합 및 반출 등에 필요한 비용을 결합신청자에게 요청할 수 있음

- (신청 방법) 결합신청자는 결합전문기관 선택 후 가명정보 결합종합지원시스템(link.privacy.go.kr)(이하 '결합종합지원시스템(link.privacy.go.kr)'이라 함)을 이용하여 결합 신청

  - 결합신청자는 결합종합지원시스템(link.privacy.go.kr)에서 결합전문기관 확인 후 선택하여 결합 신청

▶ 결합전문기관 선택에 별도의 제한은 없으므로, 결합신청자는 반출심사를 고려하여 결합 대상정보에 대한 전문성이 있는 곳을 선택하거나 분석 및 가명처리에 필요한 시스템 성능, 소요일정, 가명처리 또는 분석 지원 여부, 모의결합 지원 여부 등 결합전문기관의 지원 사항을 고려하여 결합전문기관을 선택 할 수 있음

▶ 결합전문기관(보호법)과 데이터전문기관(신용정보법)

- 신용정보법은 신용정보회사등의 정보와 결합하고자 하는 경우 데이터전문기관을 통해 결합하도록 규정(신용정보법 제17조의2 제1항)하고 있으므로, 신용정보회사 등과 결합하는 경우에는 데이터전문기관에 결합을 신청하여야 함

  ※ 결합되는 정보의 성격이 아닌 결합되는 정보를 보유한 기관에 따라 결합전문기관(보호법) 또는 데이터전문기관(신용정보법)을 구분하여 결합신청 필요

- 신용정보회사등이 아닌 기관이 보유한 금융·신용정보는 보호법에 따라 결합전문기관을 통해 결합을 수행하여야 함

○ (신청 서류) 결합신청자는 '가명정보 결합 및 반출 등에 관한 고시(이하 '결합 고시'라 함)'의 [별지 제3호] 결합신청서와 첨부 서류*를 결합 신청 시 제출하여야 함

* 단, 결합신청자가 결합의 선택사항(모의결합, 가명정보 추출 등) 진행 등의 사유로 결합대상정보에 대한 검토 및 확정이 완료되지 않은 경우, 결합전문기관과 협의하여 결합대상정보의 가명처리 내역에 관한 서류는 가명정보 전송시 제출할 수 있음

※ 결합신청서 및 첨부 서류의 구체적인 작성 방법은 「참고자료」(80p) 참고

○ (신청서 검토 및 접수) 결합전문기관이 신청서 작성내용(결합 목적 적합성 등) 및 첨부 서류에 대한 보완을 요청한 경우 결합신청자는 해당사항을 보완하여 다시 제출하여야 함

- (접수통지) 결합전문기관은 서류 누락 등 신청서류에 더 이상 보완사항이 없는 경우 결합신청서를 접수하고 결합신청자에게 신청접수 사실을 통지

- (내용보완) 결합신청이 접수된 이후 결합전문기관이 결합 목적, 결합대상 항목 등이 적절한지 여부를 추가로 확인하며, 필요한 경우 목적에 관한 서류제출이나 결합대상 변경 등을 요청할 수 있음

○ (결합 일정 및 절차 등 협의) 결합신청자는 결합 신청 내역에 따라 결합 절차 및 필요한 정보 등을 결합전문기관 및 결합키관리기관과 협의하여야 함

## 2단계 결합 수행

### 1. 결합키 생성

- 가명정보를 제공하고자 하는 결합신청자는 결합키관리기관과 결합키 생성에 관한 사항을 협의*하고 결합키관리기관으로부터 결합키 생성에 필요한 Salt값을 전송받아야 함

  * 결합키 생성 항목, 인코딩 방식, 결합키 생성 알고리즘

  - 결합신청자 간의 동일한 결합키 생성을 위해서는 결합신청자 간에 서로 협의한 바에 따라 결합키 생성 항목, 인코딩* 방법, 알고리즘을 동일하게 사용하여야 함

    * 한글의 경우, 인코딩 방식(EUC-KR, UTF-8)에 따라 동일한 일방향 암호화 알고리즘으로 데이터를 암호화하는 경우에도 서로 다른 값으로 데이터가 만들어지며 이 경우 결합이 되지 않음(UTF-8 인코딩을 권고)

**〈 결합키 생성 예시 〉**

'홍길동'+'01012345678'+'생년월일'+'abc123'
(성명/전화번호/생년월일/Salt값)


'a9fe0r … sr13'

▶ 결합키 생성 시 결합키의 대상은 일반적으로 성명, 전화번호, 생년월일 등 특정 개인을 식별할 수 있는 정보들이 사용 됨

▶ 또한, 위 정보들을 그대로 사용하면 특정 개인을 식별할 수 있으므로 결합키 생성 시에는 일반적으로 일방향 암호화 알고리즘을 사용

※ 일방향 암호화 알고리즘은 가명정보의 보호에 큰 영향을 미치게 되며 2020년 기준으로 일방향 암호화 기법 중 SHA2-512에 Salt를 포함하여 사용하거나 HMAC-SHA2 알고리즘을 이용할 것을 권고하고 있음

※ Salt의 길이는 Hash처리 결과값의 크기와 동일한 크기를 사용하는 것이 안전함

* 출처: '개인정보의 암호화 조치 안내서(2020.12.)', KISA

- 결합신청자는 결합대상정보에 정보주체별로 중복되지 않는 일련의 값(일련번호*)을 생성하여야 함

  * 모의결합 시에는 일련번호가 활용되지 않으므로, 모의결합 절차가 종료된 이후 일련번호를 생성할 수 있음

〈 일련번호 생성 예시 〉

신청기관 A

| 일련번호 | 성명 | 전화번호 | 생년월일 | ... |
|---|---|---|---|---|
| A1 | 강감찬 | 090-4562-7895 | 1947 | ... |
| A2 | 권율 | 090-7854-5689 | 1975 | ... |
| A3 | 유관순 | 090-4567-9876 | 1982 | ... |
| ... | ... | ... | ... | ... |

신청기관 B

| 일련번호 | 성명 | 전화번호 | 생년월일 | ... |
|---|---|---|---|---|
| B1 | 유관순 | 090-4567-9876 | 1982 | ... |
| B2 | 권율 | 090-7854-5689 | 1975 | ... |
| B3 | 강감찬 | 090-4562-7895 | 1947 | ... |
| ... | ... | ... | ... | ... |

- 반복결합을 신청하는 경우 추후 반출되는 정보와의 연계·분석을 위하여 결합키에 사용된 결합키 생성항목, 인코딩 방식, 알고리즘(Salt값 제외)을 보관하여야 함

## 2. 모의결합(선택사항) * 결합전문기관이 지원가능한 경우

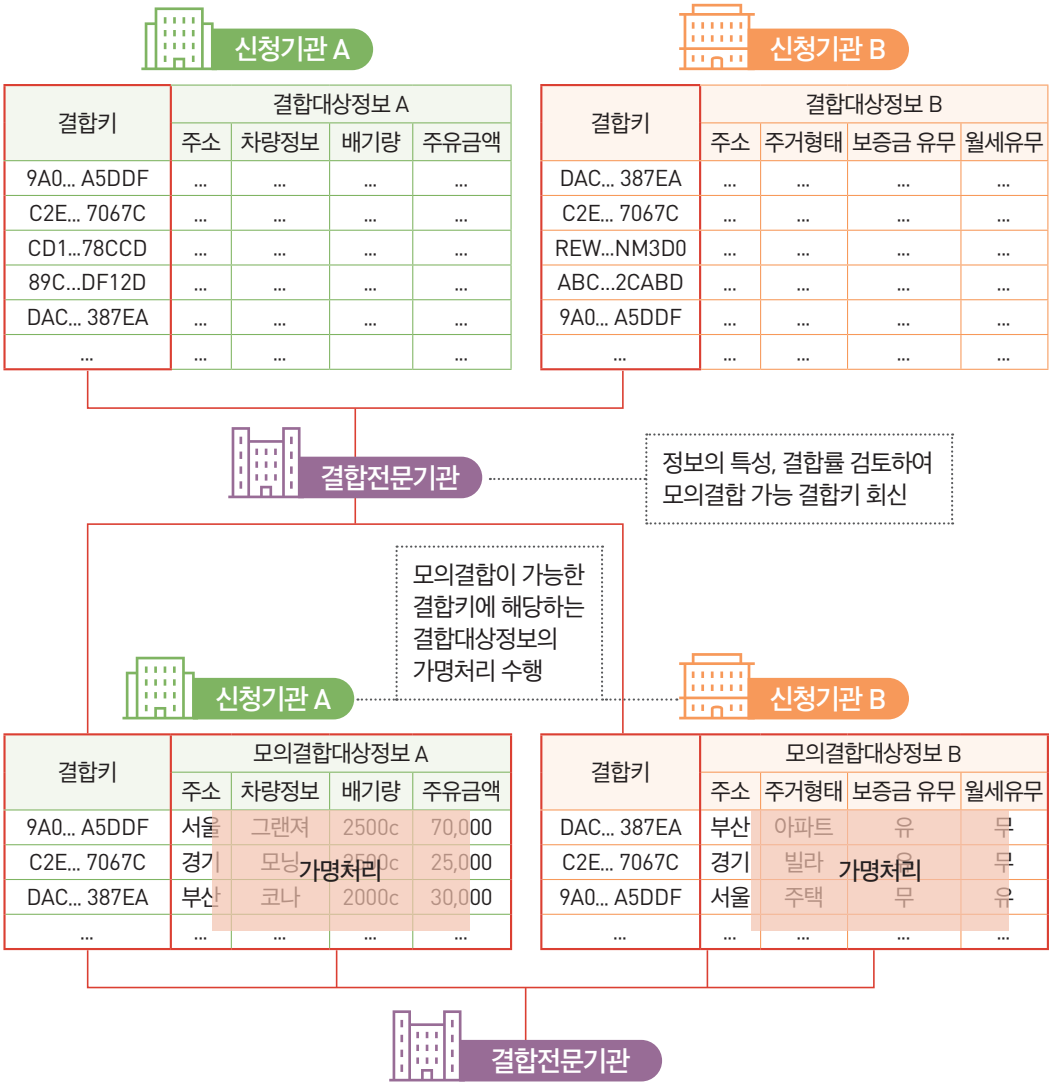
- 모의결합을 신청한 결합신청자는 결합키관리기관과의 협의에 따라 생성한 결합키를 결합전문기관에 전송하여야 함

  ※ 단, 결합신청자는 모의결합을 지원하는 결합전문기관을 확인 후 신청하여야 함

- 결합전문기관은 모의결합대상정보의 특성, 결합률 등을 고려하여 모의결합이 가능한 경우 모의결합대상정보를 선정하여 해당 결합키를 결합신청자에게 전송하여야 함

  ※ 결합전문기관은 개인정보 침해의 우려가 없는 범위 내에서 결합의 유용성 확인을 할 수 있도록 모의결합대상정보를 선정

- 결합신청자는 결합전문기관으로부터 결합키를 제공받아 해당 모의결합대상정보를 가명처리하여 가명처리 내역과 함께 결합전문기관에 전송하여야 함



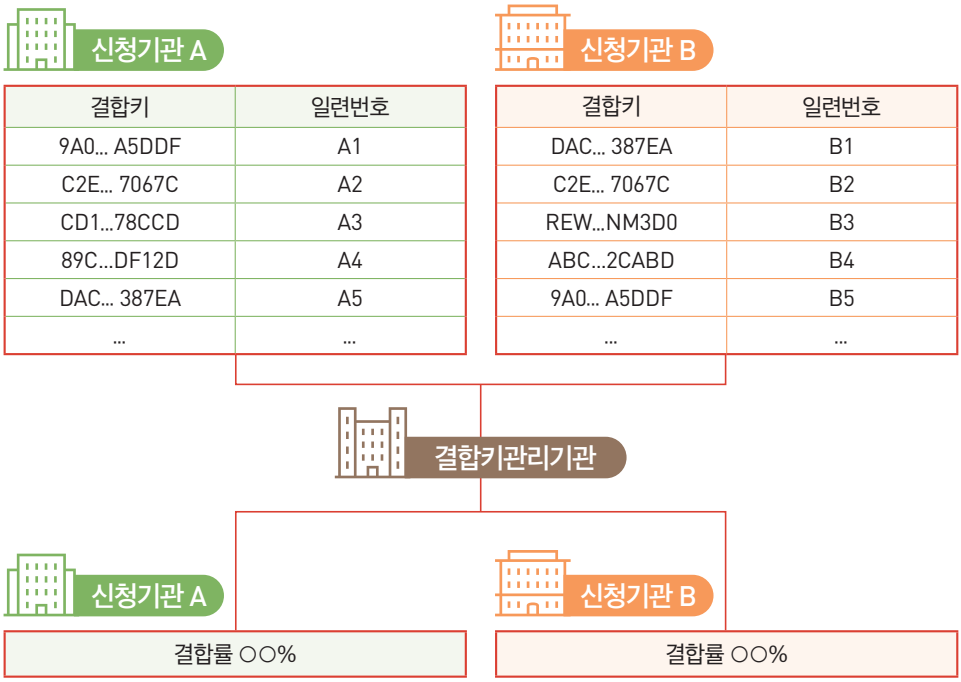

- 결합전문기관은 가명처리 내역을 확인한 후 보완 사항이 없으면 결합키를 사용하여 모의결합대상 정보의 결합을 수행함

- 결합신청자는 결합전문기관 내에서 모의결합된 정보를 분석할 수 있음
  - 단, 결합신청자는 분석한 결과물 및 모의결합된 정보를 반출할 수 없음
  ※ 결합신청자는 모의결합 분석 결과에 따라 결합의 진행 또는 종료를 결정할 수 있음

- 결합전문기관은 결합신청자의 분석이 완료된 후에는 모의결합에 사용된 정보를 파기하여야 함

## 3. 결합률 확인(선택사항)

- 결합률 확인을 신청한 결합신청자는 결합키와 일련번호를 결합키관리기관에 전송하여야 함




- 결합키관리기관은 결합신청자로부터 결합키와 일련번호를 제공받아 결합률을 확인한 후 통지하여야 함

  ※ 결합신청자는 결합률 확인 후 결합의 진행 또는 종료를 결정할 수 있음

## 4. 가명정보 추출(선택사항)

- 가명정보 추출을 신청한 결합신청자는 결합키와 일련번호를 결합키관리기관에 전송하여야 함

- 결합키관리기관은 추출 여부를 판단하는데 필요한 정보(결합 목적 등)를 결합신청자로부터 제공받아 추출 가능 여부를 검토하고, 추출이 가능한 경우 추출에 필요한 일련번호를 결합신청자에게 전송
  - 추출 가능성 검토 방법: 결합키관리기관은 결합대상정보의 특성, 결합률 등을 고려하여 추출 가능 여부를 검토
  - 추출 대상 선정 방법: 결합키관리기관은 모집단의 크기, 결합률 등을 고려하여 개인정보 침해 우려가 없도록 일정 비율의 결합되지 않는 정보(비결합대상정보)의 일련번호를 추가하여 선정

## 5. 가명처리 및 검토

- (가명처리) 결합신청자는 모의결합, 결합률 확인, 가명정보 추출 등의 선택절차가 모두 완료되고 결합절차를 진행하기로 결정하면 결합대상정보를 가명처리하여 결합전문기관에 전송하여야 함

  ※ 결합신청자는 결합전문기관에 결합대상정보의 가명처리를 지원해줄 것을 요청할 수 있음 결합 전 가명처리를 지원하는 결합전문기관을 확인 후 신청하여야 함

  - 결합신청자 중 가명정보 추출을 신청한 자는 결합키관리기관이 제공한 추출에 필요한 일련번호를 확인하고, 해당 일련번호의 결합대상정보를 가명처리한 내역(결합대상정보, 가명처리 내역, 일련번호)을 결합전문기관에 전송하여야 함

▶ 결합신청자가 보유한 개인정보파일 (예시)

- A사 : (성명, 전화번호, 생년월일), 주소, 차량 정보, 배기량, 주유금액 등
- B사 : (성명, 전화번호, 생년월일), 주소, 주거형태, 보증금 유무, 월세 유무 등

〈 결합키 생성 항목 〉

▶ 결합신청서 작성 전 처리되어야 할 가명처리 대상 항목

- A사는 주소, 차량 정보, 배기량, 주유금액 등이, B사는 주소, 주거형태, 보증금 유무, 월세 유무 등이 가명처리 대상

  * 가명처리 대상 중 분석목적에 필요하며, 식별 가능성이 현저히 낮은 항목인 경우 처리대상에서 제외 가능

  ※ 결합키 생성 항목을 결합대상정보로 활용하고자 하는 경우 식별 가능성이 존재하지 않는 것을 확인한 후 활용하여야 함

- (가명처리 검토) 결합전문기관은 결합신청자가 제출한 결합대상정보 및 가명처리 내역을 검토하여야 함
  - 만약, 보완이 필요한 경우 결합전문기관은 결합신청자에게 보완 사항을 적시하여 보완 요청

## 6. 결합

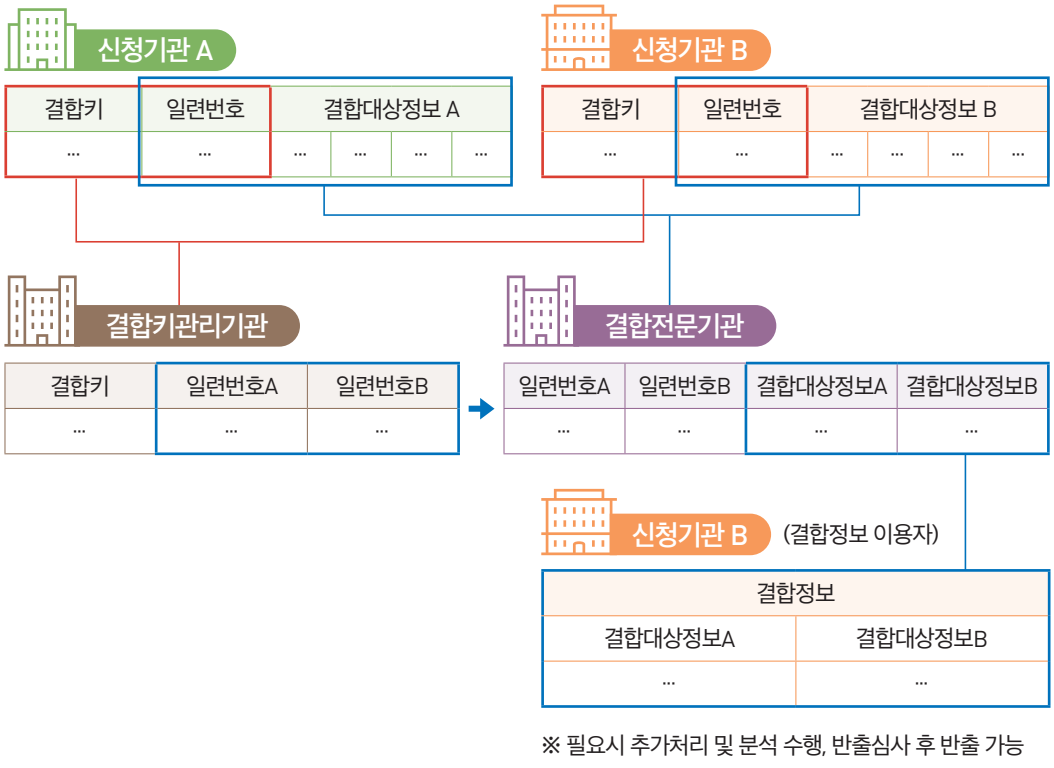
- (결합키 및 일련번호 전송) 결합신청자는 결합키와 일련번호를 결합키관리기관에 전송

  ※ 만약 결합률 확인, 가명정보 추출을 신청했고, 그 이후 결합대상정보의 변경이 없다면 이 단계에서 결합키관리기관에 정보를 전송하지 않아도 됨

- (결합키연계정보 생성) 결합키관리기관은 결합키와 일련번호를 사용하여 결합키연계정보를 생성하고 결합전문기관에 결합키연계정보 전송

  ※ 반복결합의 경우, 반복결합키를 포함한 결합키연계정보가 생성됨

- (결합) 결합전문기관은 결합키연계정보와 일련번호, 결합대상정보를 사용하여 결합




※ 필요시 추가처리 및 분석 수행, 반출심사 후 반출 가능

## 3단계 추가처리 및 반출

### 1. 추가처리 및 분석

- (추가처리) 결합정보를 이용하고자 하는 결합신청자는 결합전문기관 내에서 결합정보가 특정 개인을 알아 볼 수 있는지 여부를 확인하고, 개인식별 가능성이 확인된 경우 해당 부분에 대한 추가처리를 수행하여야 함

  ※ 결합신청자는 결합전문기관으로부터 추가처리에 대한 자문 및 처리 지원을 받을 수 있음
  단, 반출전 추가처리 지원을 지원하는 결합전문기관을 결합종합지원시스템(link.privacy.go.kr)에서 확인 후 신청하여야 함

  - 결합 이후에도 가명정보의 재식별 가능성이 증가하지 않았거나 추가 처리가 필요하지 않다고 판단되는 경우 별도의 추가처리 없이 바로 반출절차 진행 가능

- (분석) 결합신청자는 결합전문기관 내에 마련된 분석에 필요한 시설, 장비를 갖춘 공간에서 결합정보를 분석할 수 있

  ※ 결합신청자는 결합전문기관에 결합정보의 분석을 지원해줄 것을 요청할 수 있음
  단, 반출전 분석지원을 지원하는 결합전문기관을 결합종합지원시스템(link.privacy.go.kr)에서 확인 후 신청하여야 함

### 2. 반출신청 및 심사

- (반출신청) 결합정보를 반출하려는 결합신청자는 결합 고시 [별지 제4호] 반출신청서와 첨부서류*를 제출하여야 함

  * 추가적인 서류 제출이 필요한 경우에 한하여 추가 처리 내역, 반출정보를 증명할 수 있는 서류, 반출정보에 대한 안전조치 계획을 제출

  ※ 반출신청서 및 첨부 서류의 구체적인 작성 방법은 부록 (83p) 참고

- (신청서 검토 및 접수) 반출신청서를 제출받은 결합전문기관은 신청서 및 첨부 서류에 누락이 없는지 확인하고 보완 사항이 없으면 해당 반출신청서를 접수하여야 함

  - 만약, 보완할 사항이 있는 경우 결합전문기관은 해당 사유를 적시하여 결합신청자에게 보완 요청

- (반출심사) 결합신청자가 반출을 요청하면 결합전문기관은 접수일로부터 영업일 기준 5일 이내 반출심사위원회 구성 등에 관한 사항을 결합신청자에게 통지하고, 반출심사위원회를 개최하여 반출심사를 진행하여야 함
  - 결합신청자는 결합전문기관으로부터 회의개최 일정, 회의개최 장소, 반출가능 예정 시기 등이 포함된 계획서를 받을 수 있음
  - 시계열 분석을 위한 반복결합의 추가반출이면서 최초 반출과 결합대상, 가명처리 방법 등이 거의 동일한 경우, 서면회의 등으로 간소화하여 심사할 수 있음
  - (반출심사위원회 구성) 반출심사위원회는 3명의 위원으로 구성하며 반출심사를 위해 필요하다고 판단되는 경우 다른 결합전문기관에 소속된 전문가를 추가로 포함할 수 있음
  - (반출심사위원의 자격) 반출심사위원은 개인정보 보호와 관련한 업무 경력이 있거나 관련 단체로부터 추천을 받은 사람, 개인정보처리자로 구성된 단체에서 활동한 경력이 있거나 관련 단체로부터 추천을 받은 사람, 그 밖에 개인정보 보호와 관련한 경력과 전문성이 있는 사람이어야 함
  - (반출심사 기준) 반출심사는 결합 목적과 반출정보가 관련성이 있는지, 특정 개인을 알아볼 수 있지는 않은지, 반출정보에 대한 안전조치 계획이 수립되어 있는지 등을 심사하여야 함(보호법 시행령 제29조의3제4항, 결합 고시 제11조제3항)

- (추가설명 등) 결합신청자는 반출심사위원회의 요청에 따라 추가 서류를 제출하거나 직접 출석하여 설명할 수 있음

### 3. 반출

- (반출정보) 결합전문기관이 반출 승인을 하면 결합신청자는 결합전문기관 내에서 결합정보를 분석한 결과물만을 반출하거나, 결합정보(데이터셋)를 반출할 수 있음

| 1단계 | 2단계 | 3단계 | 4단계 |
|---|---|---|---|
| 사전준비 및 결합 신청 | 결합 수행 | 추가처리 및 반출 | 활용 및 사후관리 |

## 4단계 활용 및 사후관리

- (반출정보의 이용 범위) 반출정보는 결합신청자가 반출심사 시 제출한 환경(가명정보 활용 형태, 처리 장소, 방법)과 목적범위 내에서 활용하는 것이 원칙임

  ※ 결합신청자는 결합전문기관에 반출정보에 대한 분석을 지원해줄 것을 요청할 수 있음
  반출 후 분석 지원을 지원하는 결합전문기관을 결합종합지원시스템(link.privacy.go.kr)에서 확인 후 신청하여야 함

  - 결합신청자가 반출한 정보를 반출심사 시와 다른 목적으로 활용하거나 제3자에게 제공하는 것이 금지되어 있지는 않으나(보호법 제28조의2 제1항), 반출심사 시 제출한 처리 상황의 변경이 있는 경우 해당 처리 상황에 맞게 가명처리를 재수행한 후 활용하여야 함

- (반복결합) 반복결합을 신청하여 반출한 경우에는 반출정보에 반복결합키가 포함되어 있으므로, 이를 이용하여 내부에서 연계하여 활용

- (재식별 금지) 결합신청자는 반출정보를 특정 개인을 알아보기 위한 목적으로 처리하여서는 아니 되며(보호법 제28조의5 제1항), 재식별되지 않도록 지속적으로 모니터링 하여야 함

- (안전조치) 결합신청자는 반출정보를 활용하는 경우 안전성 확보에 필요한 기술적·관리적·물리적 조치를 수행하여야 함

  ※ 결합신청자는 결합전문기관에 개인정보 보호 교육을 지원해줄 것을 요청할 수 있음
  개인정보 보호 교육 지원을 지원하는 결합전문기관을
  결합종합지원시스템(link.privacy.go.kr)에서 확인 후 신청하여야 함

  ※ 안전조치에 관한 구체적인 사항은 「Ⅳ. 가명정보의 안전한 관리」(45p) 참고

# Ⅳ. 가명정보의 안전한 관리

## 1 관리적 보호조치

□ 개인정보처리자는 가명정보 또는 추가정보의 안전한 관리를 위하여 내부 관리계획의 수립, 수탁자 관리·감독 등의 관리적 보호조치를 하여야 함

**① 개인정보처리자는 가명정보 및 추가정보를 안전하게 관리하기 위한 내부 관리계획을 수립·시행하여야 함(보호법 시행령 제29조의5 제1항 제1호)**

※ 다만, 개인정보 개념에 가명정보 개념이 포함되므로, 개인정보의 안전한 관리를 위하여 수립·시행된 내부 관리계획이 있을 경우, 가명정보의 처리에 관한 내용만 추가하여 수립·시행하는 것도 가능

**제29조의5(가명정보에 대한 안전성 확보 조치)** ① 개인정보처리자는 법 제28조의4 제1항에 따라 가명정보 및 가명정보를 원래의 상태로 복원하기 위한 추가 정보(이하 이 조에서 "추가정보"라 한다)에 대하여 다음 각 호의 안전성 확보 조치를 해야 한다.

1. 제30조 또는 제48조의2에 따른 안전성 확보조치
2. 가명정보와 추가정보의 분리 보관. 다만, 추가정보가 불필요한 경우에는 추가정보를 파기해야 한다.
3. 가명정보와 추가정보에 대한 접근 권한의 분리. 다만, 「소상공인 보호 및 지원에 관한 법률」 제2조에 따른 소상공인으로서 가명정보를 취급할 자를 추가로 둘 여력이 없는 경우 등 접근 권한의 분리가 어려운 정당한 사유가 있는 경우에는 업무 수행에 필요한 최소한의 접근 권한만 부여하고 접근 권한의 보유 현황을 기록으로 보관하는 등 접근 권한을 관리·통제해야 한다.

② 법 제28조의4제2항에서 "대통령령으로 정하는 사항"이란 다음 각 호의 사항을 말한다.

1. 가명정보 처리의 목적
2. 가명처리한 개인정보의 항목
3. 가명정보의 이용내역
4. 제3자 제공 시 제공받는 자
5. 그 밖에 가명정보의 처리 내용을 관리하기 위하여 보호위원회가 필요하다고 인정하여 고시하는 사항

- 내부 관리계획에는 추가정보의 별도 분리 보관 및 이에 대한 접근권한 분리에 대한 사항 등을 포함하여야 함

〈 가명정보 처리 내부 관리계획에 포함될 사항(예시) 〉

가. 가명정보 및 추가정보의 관리책임자 지정에 관한 사항
나. 가명정보 및 추가정보의 분리 보관에 관한 사항
다. 가명정보 및 추가정보에 대한 접근권한 분리에 관한 사항
라. 가명정보 또는 추가정보의 안전성 확보조치에 관한 사항
마. 가명정보를 처리하는 자의 교육에 관한 사항
바. 가명정보 처리 기록 작성 및 보관에 관한 사항
사. 개인정보 처리방침 공개에 관한 사항
아. 가명정보의 재식별 금지에 관한 사항
※ 상기 내용에 포함되지 않은 항목은 '개인정보의 안전성 확보조치 기준 해설서' 참조

※ 가명정보 처리 내부 관리계획 작성 예시는 「참고자료」(85p) 참고

- 개인정보처리자는 내부 관리계획에서 정한 사항에 중요한 변경이 있는 경우 이를 즉시 반영하여 내부 관리계획을 수정·시행하고, 관리책임자는 연 1회 이상 내부 관리계획의 이행 실태를 점검·관리 하여야 함

② 수탁자 관리·감독의 의무(보호법 제26조)

- 개인정보처리자는 가명정보 처리업무를 외부에 위탁하는 경우, 가명정보도 개인정보에 해당하므로 보호법 제26조에 따라 위탁업무 수행 목적 외 가명정보의 처리 금지에 관한 사항 등을 포함한 문서를 작성하여야 함

- 또한, 위탁자는 위탁하는 업무의 내용과 가명정보 처리업무를 위탁받아 처리하는 자를 공개하여야하며, 업무 위탁으로 인하여 가명정보가 분실·도난·유출·위조·변조·훼손 또는 재식별 되지 아니하도록 수탁자를 교육하고, 처리현황 점검 등 수탁자가 가명정보를 안전하게 처리하는지를 감독하여야 함

〈 가명정보 처리업무 위탁계약서에 포함되어야 할 사항(예시) 〉

| 구분 | 위탁계약서에 포함되어야 할 사항 |
|---|---|
| 위탁업무 수행 목적 외 처리금지 | 가명정보를 위탁받은 범위 외로 처리하는 것을 금지하는 사항 |
| 가명정보의 안전조치 사항 | 가명정보와 추가정보의 분리 보관, 가명정보와 추가정보에 대한 접근권한 분리, 가명정보에 대한 안전조치 등에 대한 사항 |
| 위탁업무의 목적 및 범위 | 가명정보를 위탁하는 목적과 범위에 대한 사항 |
| 재위탁 제한 | 재위탁 가능한 범위에 대한 사항 |
| 관리·감독에 관한 사항 | 위탁업무와 관련하여 보유하고 있는 개인정보, 가명정보, 추가정보 등에 대한 안전성 확보조치에 관한 관리·감독사항 |
| 재식별 금지 | 가명정보를 제공받거나 처리를 위탁 받은 사업자 등은 다른 정보와 결합을 통해 재식별 시도가 금지됨을 명시 |
| 재식별 위험 발생시 통지 | 가명정보가 재식별 되었거나, 재식별 가능성이 높아지는 상황이 발생한 경우에는 가명정보 처리 중지 및 위탁자에게 통지 의무 명시 |

〈 가명정보 처리업무 위탁계약서 특수조건 반영 사례(예시) 〉

**제00조(재식별 금지)**

① ○은 △으로부터 제공받은 가명정보를 ××한 목적으로 안전하게 이용하고, 이를 이용해서 개인을 재식별하기 위한 어떠한 행위도 하여서는 아니 된다.

② ○은 △으로부터 제공받은 정보가 재식별 되거나 재식별 가능성이 현저하게 높아지는 상황이 발생하면 즉시 해당 정보의 처리를 중단하고 관련 사항을 △에게 알리며, 필요한 협조를 하여야 한다.

③ ○은 제1항에서 제2항까지의 사항을 이행하지 않아 발생하는 모든 결과에 대해 형사 및 민사상 책임을 진다.

※ 가명정보를 제공받은 기업은 "○", 제공한 기업은 "△"로 표시

### ③ 개인정보 처리방침 수립 및 공개(보호법 제30조)

- 개인정보처리자는 가명정보 처리와 관련하여 아래와 같은 내용을 개인정보 처리방침에 포함하여 공개하여야 함

※ 다만, 개인정보의 처리에 대하여 기 작성한 개인정보 처리방침이 있을 경우, 가명정보 처리에 관한 내용만 추가 가능

**〈 가명정보 활용 관련 개인정보 처리방침에 포함될 사항(예시) 〉**

1. 가명정보 처리 목적
2. 가명정보 처리 기간(선택)
3. 가명정보 제3자 제공에 관한 사항(해당되는 경우)
4. 가명정보 처리의 위탁에 관한 사항(해당되는 경우)
5. 처리하는 개인정보의 항목
6. 보호법 제28조의4(가명정보에 대한 안전조치의무 등)에 따른 가명정보의 안전성 확보 조치에 관한 사항

**〈 가명정보 활용 관련 개인정보 처리방침 반영 사례(예시) 〉**

**제00조(가명정보의 처리)**

① ○○○(개인정보처리자명)는 수집한 개인정보를 특정 개인을 알아볼 수 없도록 가명처리하여 통계작성, 과학적 연구, 공익적 기록보존 등을 위하여 처리할 수 있습니다. 가명정보 처리의 위탁 및 제3자 제공은 하지 않으며, 가명정보는 재식별 되지 않도록 분리하여 별도 저장·관리하고 가명정보의 처리 내용에 대해 기록을 작성하여 보관하는 등 필요한 기술적·관리적 보호조치를 취합니다.

| 구분 | 수집·이용 목적 | 처리항목 | 보유 및 이용기간 |
|---|---|---|---|
| △△△ 연구 | 연령대별 △△ 등 분석 | 휴대전화번호, △△일시, △△유형 | 결합데이터 분석 완료시까지 |

## 2 기술적 보호조치

□ 개인정보처리자는 가명정보 및 추가정보의 분리 보관, 접근권한 관리, 접근통제 및 접속 기록의 보관 및 점검 등의 기술적 보호조치를 하여야 함

### ① 추가정보의 분리 보관(보호법 시행령 제29조의5 제1항 제2호)

○ 개인정보처리자는 추가정보를 가명정보와 분리하여 별도로 저장·관리하고, 추가정보가 가명정보와 불법적으로 결합되어 재식별에 악용되지 않도록 접근권한을 최소화하고 접근통제를 강화하는 등 필요한 조치를 적용하여야 함

- 추가정보와 가명정보는 분리하여 보관하는 것을 원칙으로 하고, 불가피한 사유로 물리적인 분리가 어려운 경우 DB 테이블 분리 등 논리적으로 분리*하는 것도 가능함

  * 논리적으로 분리할 경우 엄격한 접근통제를 적용하여야 함

※ 추가정보의 활용 목적 달성 및 불필요한 경우에는 추가정보를 파기할 수 있으며, 이 경우 파기에 대한 기록을 작성하고 보관할 필요가 있음

### ② 접근권한의 분리(보호법 시행령 제29조의5 제1항 제3호)

○ 개인정보처리자는 가명정보 또는 추가정보에 접근할 수 있는 담당자를 가명정보 처리 업무 목적 달성에 필요한 최소한의 인원으로 엄격하게 통제하여야 하며, 접근권한도 업무에 따라 차등부여 하여야 함

※ 다만, 동일한 건에 대해서 가명정보의 적정성을 검토하는 자는 가명처리 수행하는 자, 가명정보를 활용하는 자와 분리되어야 함

- 가명정보를 취급할 자를 추가로 둘 여력이 없는 경우 등 접근권한의 분리가 어려운 정당한 사유가 있는 경우*에는 업무 수행에 필요한 최소한 접근권한 부여 및 접근권한의 보유 현황을 기록으로 보관하는 등 접근권한을 관리·통제하여야 함

  * 「소상공인 보호 및 지원에 관한 법률」 제2조에 따른 소상공인 등

- 가명정보를 처리하는 자가 가명처리를 수행하는 경우를 제외하고는 특정 개인을 알아볼 수 있는 개인정보처리시스템(가명정보처리시스템 제외)에 접근할 수 없도록 제한할 필요가 있음

- 전보 또는 퇴직 등 인사이동이 발생하여 가명정보를 처리하는 자가 변경되었을 경우 지체 없이 가명정보처리시스템의 접근권한을 변경 또는 말소하여야 함

- 가명정보처리시스템의 접근권한 부여, 변경 또는 말소에 대한 내역을 기록하고, 그 기록을 최소 3년간 보관하여야 함

- 가명정보처리시스템에 접속할 수 있는 사용자 계정을 발급하는 경우 가명정보를 처리하는 자별로 사용자 계정을 발급하여야 하며, 다른 가명정보를 처리하는 자, 추가정보를 처리하는 자, 개인정보취급자와 공유되지 않도록 하여야 함

- 가명정보를 처리하는 자가 안전한 비밀번호를 설정하여 이행할 수 있도록 비밀번호 작성규칙을 수립하여 적용하여야 함

- 가명정보에 대한 처리 권한이 있는 자만이 가명정보처리시스템에 접근할 수 있도록 계정정보 또는 비밀번호를 일정 횟수 이상 잘못 입력한 경우 접근을 제한하는 등 필요한 기술적 조치를 하여야 함

③ 가명정보 처리 관련 기록 작성·보관(보호법 시행령 제29조의5 제2항)

- 개인정보처리자는 가명정보의 처리목적, 가명처리한 개인정보 항목, 가명정보의 이용내역, 제3자 제공 시 제공받는 자를 작성하여 보관하여야 함

〈 작성방법 예시 〉

## 가명정보 처리 관리대장 (파일명 예시)

| | 구분 | 내용 | 관련 파일명 |
|---|---|---|---|
| 1 | 가명정보의 처리 목적 | | |
| 2 | 가명처리한 개인정보의 항목 | | |
| 3 | 가명정보의 이용내역 | ① 책임자:<br>② 가명정보 및 추가정보를 처리하는 자:<br>③ 가명처리 일시:<br>④ 이용방법:<br>ex) 목적외 이용, 내부이용, 외부제공, 내부 결합, 결합전문기관을 통한 결합 등 | |
| 4 | 제공받는 자 (제3자 제공시) | | |

1. 가명정보의 처리 목적을 기재(통계작성, 과학적 연구, 공익적 기록보존 등)
2. 가명처리의 대상이 된 이용 항목을 말함(예: 성별, 나이, 주소 등)
3. 가명정보 및 추가정보에 대한 책임자, 가명정보 및 추가정보를 처리하는 자(필요시 처리자 명단), 가명처리한 일시, 가명정보의 이용방법(목적외 이용, 내부이용, 외부제공, 내부 결합, 결합전문기관을 통한 결합 등)
4. (제3자에게 제공하는 경우) 가명정보를 제공받는 자의 명칭

## 3 물리적 보호조치

□ 개인정보처리자는 가명정보 또는 추가정보의 안전한 관리를 위하여 물리적 안전조치를 취하여야 함

○ 개인정보처리자는 가명정보 또는 추가정보를 전산실이나 자료보관실에 보관하는 경우 비인가자의 접근으로부터 보호하기 위하여 출입 통제 등의 절차를 수립하여야 함

○ 또한 가명정보 또는 추가정보가 보조저장매체 등에 저장되어 있는 경우 잠금장치가 있는 안전한 장소에 보관하여야 하며, 이러한 보조저장매체 등의 반·출입 통제를 위한 보안대책을 마련하여야 함

## 4 정보주체의 권리보장

□ 개인정보처리자는 보호법 제37조에 따라 정보주체가 자신의 개인정보에 대한 가명처리 정지를 요구하는 경우 이를 보장하여야 함

○ 개인정보처리자는 정보주체의 가명처리 정지를 요구 받았을 때에는 지체 없이 해당 정보주체의 개인정보 처리의 전부 또는 일부를 정지하여야함

- 다만, 이미 해당주체의 개인정보가 가명처리된 경우에는 가명처리 정지 요구가 적용되지 않으며, 해당 정보주체의 개인정보에 대해서는 향후 통계작성, 과학적 연구, 공익적 기록보존 등 목적으로 가명처리가 이루어지지 않도록 처리하여야 함

※ 가명정보는 특정 개인을 알아볼 수 없는 정보로 현행법상 재식별이 불가하며, 이에 따라 해당 정보주체의 개인정보가 가명처리 되었는지 여부를 확인할 수 없음(보호법 제28조의5 제1항)

**제28조의5 (가명정보 처리 시 금지의무 등)** ① 누구든지 특정 개인을 알아보기 위한 목적으로 가명정보를 처리해서는 아니 된다.

# 가명정보 처리 가이드라인_ 부록

참고자료

# 참고자료

## 참고1 개인정보 가명처리 기술 및 예시

### ◻ 개인정보의 가명·익명처리 기술 종류

※ 아래 분류는 이해를 돕기 위해 ISO/IEC 20889, 그리고 EU ENISA에서 발간한 보고서 등 국내·외 자료들을 참고하여 작성했으며 표준이 아님

| 분류 | 기술 | 세부기술 | 설명 |
|---|---|---|---|
| 개인정보 삭제 | 삭제기술 | 삭제 (Suppression) | • 원본정보에서 개인정보를 단순 삭제 |
| | | 부분삭제 (Partial suppression) | • 개인정보 전체를 삭제하는 방식이 아니라 일부를 삭제 |
| | | 행 항목 삭제 (Record suppression) | • 다른 정보와 뚜렷하게 구별되는 행 항목을 삭제 |
| | | 로컬 삭제 (Local suppression) | • 특이정보를 해당 행 항목에서 삭제 |
| 개인정보 일부 또는 전부 대체 | 삭제기술 | 마스킹 (Masking) | • 특정 항목의 일부 또는 전부를 공백 또는 문자(‘*’, ‘_’ 등이나 전각 기호)로 대체 |
| | 통계도구 | 총계처리 (Aggregation) | • 평균값, 최댓값, 최솟값, 최빈값, 중간값 등으로 처리 |
| | | 부분총계 (Micro aggregation) | • 정보집합물 내 하나 또는 그 이상의 행 항목에 해당하는 특정 열 항목을 총계처리. 즉, 다른 정보에 비하여 오차 범위가 큰 항목을 평균값 등으로 대체 |
| | 일반화 (범주화) 기술 | 일반 라운딩 (Rounding) | • 올림, 내림, 반올림 등의 기준을 적용하여 집계 처리하는 방법으로, 일반적으로 세세한 정보보다는 전체 통계정보가 필요한 경우 많이 사용 |
| | | 랜덤 라운딩 (Random rounding) | • 수치 데이터를 임의의 수인 자리 수, 실제 수 기준으로 올림(round up) 또는 내림(round down)하는 기법 |
| | | 제어 라운딩 (Controlled rounding) | • 라운딩 적용 시 값의 변경에 따라 행이나 열의 합이 원본의 행이나 열의 합과 일치하지 않는 단점을 해결하기 위해 원본과 결과가 동일하도록 라운딩을 적용하는 기법 |
| | | 상하단코딩 (Top and bottom coding) | • 정규분포의 특성을 가진 데이터에서 양쪽 끝에 치우친 정보는 적은 수의 분포를 가지게 되어 식별성을 가질 수 있음<br>• 이를 해결하기 위해 적은 수의 분포를 가진 양 끝단의 정보를 범주화 등의 기법을 적용하여 식별성을 낮추는 기법 |

1) EU ENISA(European Union Agency for Network and Information Security), Recommendations on shaping technology according to GDPR provisions, An overview on data pseudonymisatin, November 2018

EU ENISA(European Union Agency for Network and Information Security), Pseudonymisation and best practices, November 2019

| 분류 | 기술 | 세부기술 | 설명 |
|---|---|---|---|
| 개인정보 일부 또는 전부 대체 | 일반화 (범주화) 기술 | 로컬 일반화 (Local generalization) | • 전체 정보집합물 중 특정 열 항목(들)에서 특이한 값을 가지거나 분포상 의 특이성으로 인해 식별성이 높아지는 경우 해당 부분만 일반화를 적용하여 식별성을 낮추는 기법 |
| | | 범위 방법 (Data range) | • 수치 데이터를 임의의 수 기준의 범위(range)로 설정하는 기법으로, 해당 값의 범위 또는 구간(interval)으로 표현 |
| | | 문자데이터 범주화 (Categorization of character data) | • 문자로 저장된 정보에 대해 보다 상위의 개념으로 범주화하는 기법 |
| | 암호화 | 양방향 암호화 (Two-way encryption) | • 특정 정보에 대해 암호화와 암호화된 정보에 대한 복호화가 가능한 암호화 기법<br>• 암호화 및 복호화에 동일 비밀키로 암호화하는 대칭키(Symmetric key) 방식과 공개키와 개인키를 이용하는 비대칭키(Asymmetric key) 방식으로 구분 |
| | | 일방향 암호화 - 암호학적 해시함수 (One-way encryption - Cryptographic hash function) | • 원문에 대한 암호화의 적용만 가능하고 암호문에 대한 복호화 적용이 불가능한 암호화 기법<br>• 키가 없는 해시함수(MDC, Message Digest Code), 솔트(Salt)가 있는 해시함수, 키가 있는 해시함수(MAC, Message Authentication Code)로 구분<br>• 암호화(해시처리)된 값에 대한 복호화가 불가능하고, 동일한 해시 값과 매핑(mapping)되는 2개의 고유한 서로 다른 입력 값을 찾는 것이 계산상 불가능하여 충돌 가능성이 매우 적음 |
| | | 순서보존 암호화 (Order-preserving encryption) | • 원본정보의 순서와 암호값의 순서가 동일하게 유지되는 암호화 방식<br>• 암호화된 상태에서도 원본정보의 순서가 유지되어 값들 간의 크기에 대한 비교 분석이 필요한 경우 안전한 분석이 가능 |
| | | 형태보존 암호화 (Format-preserving encryption) | • 원본 정보의 형태와 암호화된 값의 형태가 동일하게 유지되는 암호화 방식<br>• 원본 정보와 동일한 크기와 구성 형태를 가지기 때문에 일반적인 암호화가 가지고 있는 저장 공간의 스키마 변경 이슈가 없어 저장 공간의 비용 증가를 해결할 수 있음<br>• 암호화로 인해 발생하는 시스템의 수정이 거의 발생하지 않아 토큰화, 신용카드 번호의 암호화 등에서 기존 시스템의 변경 없이 암호화를 적용할 때 사용 |
| | | 동형 암호화 (Homomorphic encryption) | • 암호화된 상태에서의 연산이 가능한 암호화 방식으로 원래의 값을 암호화한 상태로 연산 처리를 하여 다양한 분석에 이용가능<br>• 암호화된 상태의 연산값을 복호화 하면 원래의 값을 연산한 것과 동일한 결과를 얻을 수 있는 4세대 암호화 기법 |
| | | 다형성 암호화 (Polymorphic encryption) | • 가명정보의 부정한 결합을 차단하기 위해 각 도메인별로 서로 다른 가명처리 방법을 사용하여 정보를 제공하는 방법<br>• 정보 제공 시 서로 다른 방식의 암호화된 가명처리를 적용함에 따라 도메인별로 다른 가명정보를 가지게 됨 |
| | 무작위화 기술 | 잡음 추가 (Noise addition) | • 개인정보에 임의의 숫자 등 잡음을 추가(더하기 또는 곱하기)하는 방법 |

| 분류 | 기술 | 세부기술 | 설명 |
|---|---|---|---|
| 개인정보 일부 또는 전부 대체 | 무작위화 기술 | 순열(치환) (Permutation) | • 분석 시 가치가 적고 식별성이 높은 열 항목에 대해 대상 열 항목의 모든 값을 열 항목 내에서 무작위로 순서를 변경하여 식별성을 낮추는 기법<br>• 개인정보를 다른 행 항목의 정보와 무작위로 순서를 변경하여 전체정보에 대한 변경 없이 특정 정보가 해당 개인과 연결되지 않도록 하는 방법 |
| | | 토큰화 (Tokenisation) | • 개인을 식별할 수 있는 정보를 토큰으로 변환 후 대체함으로써 개인정보를 직접 사용하여 발생하는 식별 위험을 제거하여 개인정보를 보호하는 기술<br>• 토큰 생성 시 적용하는 기술은 의사난수생성 기법이나 양방향 암호화, 형태보존 암호화 기법을 주로 사용 |
| | | (의사)난수생성기 ((P)RNG, (Pseudo) Random Number Generator) | • 주어진 입력값에 대해 예측이 불가능하고 패턴이 없는 값을 생성하는 메커니즘으로 임의의 숫자를 개인정보와 대체 |
| 가명·익명처리를 위한 다양한 기술 (기타 기술) | | 표본추출 (Sampling) | • 데이터 주체별로 전체 모집단이 아닌 표본에 대해 무작위 레코드 추출 등의 기법을 통해 모집단의 일부를 분석하여 전체에 대한 분석을 대신하는 기법 |
| | | 해부화 (Anatomization) | • 기존 하나의 데이터셋(테이블)을 식별성이 있는 정보집합물과 식별성이 없는 정보집합물로 구성된 2개의 데이터셋으로 분리하는 기술 |
| | | 재현데이터 (Synthetic data) | • 원본과 최대한 유사한 통계적 성질을 보이는 가상의 데이터를 생성하기 위해 개인정보의 특성을 분석하여 새로운 데이터를 생성하는 기법 |
| | | 동형비밀분산 (Homomorphic secret sharing) | • 식별정보 또는 기타 식별가능정보를 메시지 공유 알고리즘에 의해 생성된 두 개 이상의 쉐어(share)*로 대체<br>*기밀사항을 재구성하는데 사용할 수 있는 하위 집합 |
| | | 차분 프라이버시 (Differential privacy) | • 특정 개인에 대한 사전지식이 있는 상태에서 데이터베이스 질의(Query)에 대한 응답 값으로 개인을 알 수 없도록 응답 값에 임의의 숫자 잡음(Noise)을 추가하여 특정 개인의 존재 여부를 알 수 없도록 하는 기법<br>• 1개 항목이 차이나는 두 데이터베이스간의 차이(확률분포)를 기준으로 하는 프라이버시 보호 모델 |

## ■ 개인정보의 가명·익명처리 예시

※ 아래 모든 예시는 각 기법의 적용에 대한 예시이며 전체 데이터에 대한 가명·익명처리에 대한 예시가 아닙니다.

### ① 개인정보 삭제

▶ 삭제기술: 선택된 항목을 제거하는 기술

❶ 삭제(Suppression) 수치형데이터 문자형데이터

- 원본정보에서 개인정보를 단순 삭제

※ 이때 남아 있는 정보 그 자체로도 분석의 유효성을 가져야 함과 동시에 개인을 식별할 수 없어야 하며, 인터넷 등에 공개되어 있는 정보 등과 결합하였을 경우에도 개인을 식별할 수 없어야 함

| 성명 | 성별 | 나이 | 핸드폰번호 | 주소 | 통신료 | 단말기금액 | 누적 포인트 |
|---|---|---|---|---|---|---|---|
| 김철수 | 남 | 41세 | 010-6666-8888 | 서울특별시 중구 무교동 | 98,700 | 1,198,700 | 356,800 |
| 이영희 | 여 | 61세 | 010-9999-2222 | 부산광역시 북구 화명동 | 69,400 | 505,400 | 203,000 |
| 박민호 | 남 | 30세 | 010-2222-7777 | 광주광역시 서구 금호동 | 104,400 | 1,604,400 | 198,000 |
| 이윤정 | 여 | 57세 | 010-3333-4444 | 전라남도 나주시 빛가람동 | 954,800 | 3,954,800 | 20,532,000 |
| 최동욱 | 남 | 28세 | 010-5555-6666 | 세종특별자치시 어진동 | 83,600 | 883,600 | 400,900 |

삭제

| 성별 | 나이 | 통신료 | 단말기금액 | 누적포인트 |
|---|---|---|---|---|
| 남 | 41세 | 98,700 | 1,198,700 | 356,800 |
| 여 | 61세 | 69,400 | 505,400 | 203,000 |
| 남 | 30세 | 104,400 | 1,604,400 | 198,000 |
| 여 | 57세 | 954,800 | 3,954,800 | 20,532,000 |
| 남 | 28세 | 83,600 | 883,600 | 400,900 |

❷ 부분삭제(Partial suppression) 수치형데이터 문자형데이터

- 개인정보 전체를 삭제하는 방식이 아니라 일부를 삭제

| 성명 | 성별 | 나이 | 핸드폰번호 | 주소 | 통신료 | 단말기금액 | 누적포인트 |
|---|---|---|---|---|---|---|---|
| 김철수 | 남 | 41세 | 010-6666-8888 | 서울특별시 중구 무교동 | 98,700 | 1,198,700 | 356,800 |
| 이영희 | 여 | 61세 | 010-9999-2222 | 부산광역시 북구 화명동 | 69,400 | 505,400 | 203,000 |
| 박민호 | 남 | 30세 | 010-2222-7777 | 광주광역시 서구 금호동 | 104,400 | 1,604,400 | 198,000 |
| 이윤정 | 여 | 57세 | 010-3333-4444 | 전라남도 나주시 빛가람동 | 954,800 | 3,954,800 | 20,532,000 |
| 최동욱 | 남 | 28세 | 010-5555-6666 | 세종특별자치시 어진동 | 83,600 | 883,600 | 400,900 |

삭제

| 성명 | 성별 | 나이 | 핸드폰번호 | 주소 | 통신료 | 단말기금액 | 누적포인트 |
|---|---|---|---|---|---|---|---|
| 김 | 남 | 41세 | 8888 | 서울특별시 중구 | 98,700 | 1,198,700 | 356,800 |
| 이 | 여 | 61세 | 2222 | 부산광역시 북구 | 69,400 | 505,400 | 203,000 |
| 박 | 남 | 30세 | 7777 | 광주광역시 서구 | 104,400 | 1,604,400 | 198,000 |
| 이 | 여 | 57세 | 4444 | 전라남도 나주시 | 954,800 | 3,954,800 | 20,532,000 |
| 최 | 남 | 28세 | 6666 | 세종특별자치시 | 83,600 | 883,600 | 400,900 |

❸ 행 항목 삭제(Record suppression) 수치형데이터 문자형데이터

- 다른 정보와 뚜렷하게 구별되는 행 항목을 삭제
- 통계분석에 있어서 전체 평균에 비하여 오차범위를 벗어나는 자료를 제거할 때 사용

| 성명 | 성별 | 나이 | 핸드폰번호 | 주소 | 통신료 | 단말기금액 | 누적포인트 |
|---|---|---|---|---|---|---|---|
| 김철수 | 남 | 41세 | 010-6666-8888 | 서울특별시 중구 무교동 | 98,700 | 1,198,700 | 356,800 |
| 이영희 | 여 | 61세 | 010-9999-2222 | 부산광역시 북구 화명동 | 69,400 | 505,400 | 203,000 |
| 박민호 | 남 | 30세 | 010-2222-7777 | 광주광역시 서구 금호동 | 104,400 | 1,604,400 | 198,000 |
| 이윤정 | 여 | 57세 | 010-3333-4444 | 전라남도 나주시 빛가람동 | 954,800 | 3,954,800 | 20,532,000 |
| 최동욱 | 남 | 28세 | 010-5555-6666 | 세종특별자치시 어진동 | 83,600 | 883,600 | 400,900 |

삭제

| 성명 | 성별 | 나이 | 핸드폰번호 | 주소 | 통신료 | 단말기금액 | 누적포인트 |
|---|---|---|---|---|---|---|---|
| 김철수 | 남 | 41세 | 010-6666-8888 | 서울특별시 중구 무교동 | 98,700 | 1,198,700 | 356,800 |
| 이영희 | 여 | 61세 | 010-9999-2222 | 부산광역시 북구 화명동 | 69,400 | 505,400 | 203,000 |
| 박민호 | 남 | 30세 | 010-2222-7777 | 광주광역시 서구 금호동 | 104,400 | 1,604,400 | 198,000 |
| | | | | | | | |
| 최동욱 | 남 | 28세 | 010-5555-6666 | 세종특별자치시 어진동 | 83,600 | 883,600 | 400,900 |

### ❹ 로컬 삭제(Local suppression) 수치형데이터 문자형데이터

- 특이정보를 해당 행 항목에서 삭제

(설명) 다른 누적포인트에 비하여 뚜렷이 구별되는 누적포인트를 항목에서 삭제

| 성명 | 성별 | 나이 | 핸드폰번호 | 주소 | 통신료 | 단말기금액 | 누적 포인트 |
|---|---|---|---|---|---|---|---|
| 김철수 | 남 | 41세 | 010-6666-8888 | 서울특별시 중구 무교동 | 98,700 | 1,198,700 | 356,800 |
| 이영희 | 여 | 61세 | 010-9999-2222 | 부산광역시 북구 화명동 | 69,400 | 505,400 | 203,000 |
| 박민호 | 남 | 30세 | 010-2222-7777 | 광주광역시 서구 금호동 | 104,400 | 1,604,400 | 198,000 |
| 이윤정 | 여 | 57세 | 010-3333-4444 | 전라남도 나주시 빛가람동 | 954,800 | 3,954,800 | 20,532,000 |
| 최동욱 | 남 | 28세 | 010-5555-6666 | 세종특별자치시 어진동 | 83,600 | 883,600 | 400,900 |

삭제

| 성명 | 성별 | 나이 | 핸드폰번호 | 주소 | 통신료 | 단말기금액 | 누적 포인트 |
|---|---|---|---|---|---|---|---|
| 김철수 | 남 | 41세 | 010-6666-8888 | 서울특별시 중구 무교동 | 98,700 | 1,198,700 | 356,800 |
| 이영희 | 여 | 61세 | 010-9999-2222 | 부산광역시 북구 화명동 | 69,400 | 505,400 | 203,000 |
| 박민호 | 남 | 30세 | 010-2222-7777 | 광주광역시 서구 금호동 | 104,400 | 1,604,400 | 198,000 |
| 이윤정 | 여 | 57세 | 010-3333-4444 | 전라남도 나주시 빛가람동 | 954,800 | 3,954,800 | |
| 최동욱 | 남 | 28세 | 010-5555-6666 | 세종특별자치시 어진동 | 83,600 | 883,600 | 400,900 |

### ❺ 마스킹(Masking) 수치형데이터 문자형데이터

- 특정 항목의 일부 또는 전부를 공백 또는 문자(‘ * ’, ‘ _ ’ 등이나 전각 기호)로 대체

※ 분류는 개인정보 일부 또는 전부 대체로 분류되지만, 기술적으로 마스킹된 부분은 데이터로써의 가치가 없어져 일부 문건에서는 삭제로 분류되기도 함

| 성명 | 성별 | 나이 | 핸드폰번호 |
|---|---|---|---|
| 김철수 | 남 | 41세 | 010-6666-8888 |
| 이영희 | 여 | 61세 | 010-9999-2222 |
| 박민호 | 남 | 30세 | 010-2222-7777 |
| 이윤정 | 여 | 57세 | 010-3333-4444 |
| 최동욱 | 남 | 28세 | 010-5555-6666 |

마스킹

| 성명 | 성별 | 나이 | 핸드폰번호 |
|---|---|---|---|
| 김** | 남 | 4*세 | ***-****-**** |
| 이** | 여 | 6*세 | ***-****-**** |
| 박** | 남 | 3*세 | ***-****-**** |
| 이** | 여 | 5*세 | ***-****-**** |
| 최** | 남 | 2*세 | ***-****-**** |

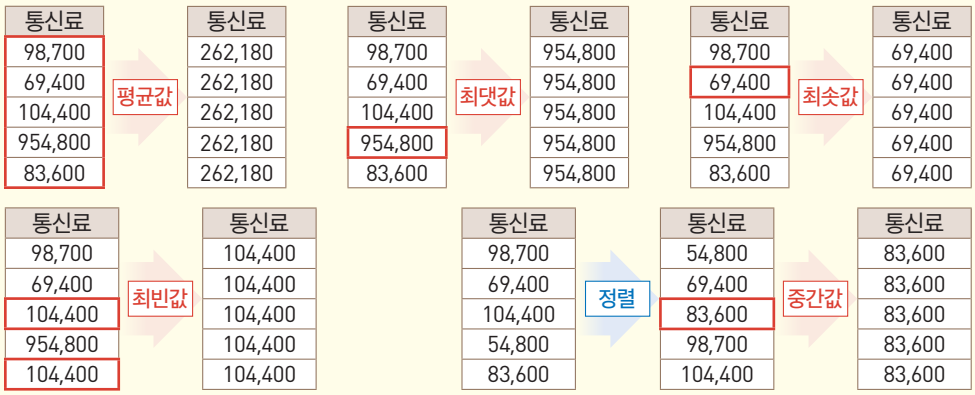
## ② 개인정보 일부 또는 전부 대체

▶ 통계도구: 데이터의 전체 구조를 변경하는 통계적 성질을 가진 기법

### ❶ 총계처리(Aggregation) 수치형데이터

- 평균값, 최댓값, 최솟값, 최빈값, 중간값 등으로 처리

※ 단, 데이터 전체가 유사한 특징을 가진 개인으로 구성되어 있을 경우 그 데이터의 대푯값이 특정 개인의 정보를 그대로 노출시킬 수도 있으므로 주의 필요




### 1-1. 부분총계(Micro Aggregation) 수치형데이터

- 정보집합물 내 하나 또는 그 이상의 행 항목에 해당하는 특정 열 항목을 총계처리즉, 다른 정보에 비하여 오차 범위가 큰 항목을 평균값 등으로 대체
- 동질 집합 내의 특정 항목을 총계처리 하거나 특정 조건에 너무 특이한 값이 있어 개인의 식별 가능성이 높지만 분석에 꼭 필요한 값인 경우 처리

(설명) 지역, 나이 기준으로 동질집합을 형성하고,
오차 범위가 큰 소득금액을 동질집합 내 평균값으로 대체

| 지역 | 나이 | 소득금액 |
|---|---|---|
| 서울 | 30대 | 5,987,900 |
| 서울 | 30대 | 28,169,700 |
| 서울 | 30대 | 3,009,600 |
| 나주 | 30대 | 4,607,300 |
| 나주 | 30대 | 3,560,800 |
| 나주 | 30대 | 2,940,100 |
| 세종 | 30대 | 6,088,400 |
| 세종 | 30대 | 2,789,200 |
| 세종 | 30대 | 5,048,300 |

| 지역 | 나이 | 소득금액 |
|---|---|---|
| 서울 | 30대 | 12,389,067 |
| 서울 | 30대 | 12,389,067 |
| 서울 | 30대 | 12,389,067 |
| 나주 | 30대 | 4,607,300 |
| 나주 | 30대 | 3,560,800 |
| 나주 | 30대 | 2,940,100 |
| 세종 | 30대 | 6,088,400 |
| 세종 | 30대 | 2,789,200 |
| 세종 | 30대 | 5,048,300 |

▶ 일반화기술: 범주화로도 불리며, 특정한 값을 상위의 속성으로 대체

### ❶ 라운딩(Rounding) 수치형데이터

#### 1-1. 일반 라운딩

- 올림, 내림, 반올림 등의 기준을 적용하여 집계 처리하는 방법

| 나이 |
|---|
| 33세 |
| 61세 |
| 47세 |
| 66세 |
| 40세 |

| 올림 | 내림 | 반올림 |
|---|---|---|
| 40세 | 30세 | 30세 |
| 70세 | 60세 | 60세 |
| 50세 | 40세 | 50세 |
| 70세 | 60세 | 70세 |
| 40세 | 40세 | 40세 |

※ 적절하지 않은 라운딩의 경우 라운딩 후에도 남은 값의 유일성이 남게 될 수 있으며, 적용하는 단위에 대한 판단이 중요

| 금액 | 백 단위 라운딩 |
|---|---|
| 983,116,785 | 983,117,000 |
| 984,715,591 | 984,716,000 |
| 984,932,383 | 984.932.000 |
| 985,660,262 | 985,660,000 |
| 986,047,778 | 986,048,000 |

적절하지 않은 라운딩

| 금액 | 백만 단위 라운딩 |
|---|---|
| 983,116,785 | 980,000,000 |
| 984,715,591 | 980,000,000 |
| 984,932,383 | 980,000,000 |
| 985,660,262 | 990,000,000 |
| 986,047,778 | 990,000,000 |

적절한 라운딩

#### 1-2. 랜덤 라운딩(Random Rounding) 수치형데이터

- 수치 데이터를 임의의 수인 자리 수, 실제 수 기준으로 올림(round up) 또는 내림(round down)하는 기법

| 금액 | | 금액 |
|---|---|---|
| 869,250 | 만 단위 라운딩 | 900,000 |
| 4,559,120 | 십만 단위 라운딩 | 4,000,000 |
| 13,601,564 | 십만 단위 라운딩 | 14,000,000 |
| 979,118 | 만 단위 라운딩 | 900,000 |
| 122,848,878 | 백만 단위 라운딩 | 120,000,000 |

### 1-3. 제어 라운딩(Controlled rounding) 수치형데이터

- 라운딩 적용 시 값의 변경에 따라 행이나 열의 합이 원본의 행이나 열의 합과 일치하지 않는 단점을 해결하기 위해 원본과 결과가 동일하도록 라운딩을 적용하는 기법

※ 컴퓨터 프로그램으로 구현하기 어렵고 복잡한 통계표에는 적용하기 어려우며, 해결할 수 있는 방법이 존재하지 않을 수 있어 아직 실무에서는 잘 사용하지 않음

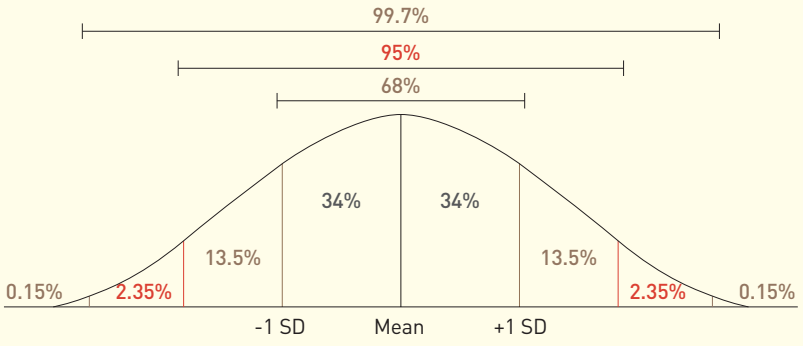
(설명) 나이에 대한 평균 분석 시 원본의 경우 평균이 51세가 되나 일반 라운딩을 적용한 경우 평균이 50세가 되어 결과가 다르게 되고, 이에 일부 값을 다르게 라운딩(제어)하여 평균 나이가 원본과 일치되도록 함

| 원본(나이) |
|---|
| 33세 |
| 61세 |
| 50세 |
| 72세 |
| 43세 |
| 44세 |
| 23세 |
| 67세 |
| 68세 |
| 49세 |
| 평균: 51세 |
| 합계: 510 |

| 일반 라운딩 | 제어 라운딩 |
|---|---|
| 30세 | 30세 |
| 60세 | 60세 |
| 50세 | 50세 |
| 70세 | 70세 |
| 40세 | 40세 |
| 40세 | 50세 |
| 20세 | 20세 |
| 70세 | 70세 |
| 70세 | 70세 |
| 50세 | 50세 |
| 평균: 50세 | 평균: 51세 |
| 합계: 500 | 합계: 510 |

### ❷ 상하단코딩(Top and bottom coding) 수치형데이터

- 정규분포의 특성을 가진 데이터에서 양쪽 끝에 치우친 정보는 적은 수의 분포를 가지게 되어 식별성을 가질 수 있으며, 이를 해결하기 위해 적은 수의 분포를 가진 양 끝단의 정보를 범주화 등의 기법을 적용하여 식별성을 낮추는 기법



❸ 로컬 일반화(Local generalization) 수치형데이터

- 전체 정보집합물 중 특정 열 항목(들)에서 특이한 값을 가지거나 분포상의 특이성으로 인해 식별성이 높아지는 경우 해당 부분만 일반화를 적용하여 식별성을 낮추는 기법

(설명) 서울 지역의 30대 중 분포 상 다른 금액에 비해 특이한 값을 동질집합 내 범주화
※ 특이한 로컬(28,169,700)에만 3,009,600 ~ 28,169,700으로 범주화 할 수 있음

| 지역 | 나이 | 소득금액 |
|---|---|---|
| 서울 | 30대 | 5,987,900 |
| 서울 | 30대 | 28,169,700 |
| 서울 | 30대 | 3,009,600 |
| 나주 | 30대 | 4,607,300 |
| 나주 | 30대 | 3,560,800 |
| 나주 | 30대 | 2,940,100 |
| 세종 | 30대 | 6,088,400 |
| 세종 | 30대 | 2,789,200 |
| 세종 | 30대 | 5,048,300 |

| 지역 | 나이 | 소득금액 |
|---|---|---|
| 서울 | 30대 | 3,009,600~28,169,700 |
| 서울 | 30대 | 3,009,600~28,169,700 |
| 서울 | 30대 | 3,009,600~28,169,700 |
| 나주 | 30대 | 4,607,300 |
| 나주 | 30대 | 3,560,800 |
| 나주 | 30대 | 2,940,100 |
| 세종 | 30대 | 6,088,400 |
| 세종 | 30대 | 2,789,200 |
| 세종 | 30대 | 5,048,300 |

❹ 범위 방법(Data range) 수치형데이터

- 수치 데이터를 임의의 수 기준의 범위(range)로 설정하는 기법으로, 해당 값의 범위 또는 구간(interval)으로 표현

(예시) 소득 3,300만원을 소득 3,000만원~4,000만원으로 대체 표기

❺ 문자데이터 범주화(Categorization of character data) 문자형데이터

- 문자로 저장된 정보에 대해 상위의 개념으로 범주화하는 기법

| 품목 | 품목 |
|---|---|
| 분유 | 육아용품 |
| 기저귀 | 육아용품 |
| 젖병 | 육아용품 |
| 샤워타올 | 육아용품 |
| 욕실화 | 육아용품 |

▶ 암호화: 정보 가공 시 일정한 규칙의 알고리즘을 적용하여 대체

❶ 암호화(Encryption) 수치형데이터 문자형데이터

※ 암호화에 따른 세부적인 내용은 한국인터넷진흥원 암호이용활성화 관련 안내서 참조

1-1. 양방향 암호화(Two-way encryption)

- 특정 정보에 대해 암호화와 암호화된 정보에 대한 복호화가 가능한 암호화 기법
- 암호화 및 복호화에 동일한 비밀키로 암호화하는 AES, ARIA 등 대칭키(Symmetric key) 방식과 공개키와 개인키를 이용하는 RSA 등 비대칭키(Asymmetric key) 방식으로 구분되며, 키(key) 관리에 주의 필요

1-2. 일방향 암호화 - 암호학적 해시함수(One-way encryption - Cryptographic hash function)

- 원문에 대한 암호화의 적용만 가능하고 암호문에 대한 복호화 적용이 불가능한 암호화 기법
- 키가 없는 해시함수(MDC, Message Digest Code), 키가 있는 해시함수(MAC, Message Authentication Code), 솔트(Salt)가 있는 해시함수로 구분
- 암호화(해시처리)된 값에 대한 복호화가 불가능하고, 동일한 해시 값과 매핑(mapping)되는 2개의 고유한 서로 다른 입력값을 찾는 것이 계산상 불가능하여 충돌 가능성이 매우 적음

1-3. 순서보존 암호화(Order-preserving encryption)

- 원본정보의 순서와 암호값의 순서가 동일하게 유지되는 암호화 방식
- 암호화된 상태에서도 원본정보의 순서가 유지되어 값들 간의 크기에 대한 비교 분석이 필요한 경우 안전한 분석이 가능

1-4. 형태보존 암호화(Format-preserving encryption)

- 원본 정보의 형태와 암호화된 암호값의 형태가 동일하게 유지되는 암호화 방식
- 원본 정보와 동일한 크기와 구성 형태를 가지기 때문에 일반적인 암호화가 가지고 있는 저장 공간의 스키마 변경 이슈가 없어 저장 공간의 비용 증가를 해결할 수 있음
- 암호화로 인해 발생하는 시스템의 수정이 거의 발생하지 않아 토큰화, 신용카드 번호의 암호화 등에서 기존 시스템의 변경 없이 암호화를 적용할 때 사용

1-5. 동형 암호화(Homomorphic encryption)

- 암호화된 상태에서의 연산이 가능한 암호화 방식
- 원래의 값을 암호화한 상태로 연산 처리를 하여 다양한 분석에 이용가능
- 암호화된 상태의 연산한 값을 복호화 하면 원래의 값을 연산한 것과 동일한 결과를 얻을 수 있는 4세대 암호화 기법

1-6. 다형성 암호화(Polymorphic encryption)

- 가명정보의 부정한 결합을 차단하기 위해 각 도메인별로 서로 다른 가명처리 방법을 사용하여 정보를 제공하는 방법
- 정보 제공 시 서로 다른 방식의 암호화된 가명처리를 적용함에 따라 도메인별로 다른 가명정보를 가지게 됨

▶ 무작위화기술: 속성의 값을 원래의 값과 다르게 변경

### ❶ 잡음 추가(Noise addition) 수치형데이터 문자형데이터

- 개인정보에 임의의 숫자 등 잡음을 추가(더하기 또는 곱하기)하는 방법
- 지정된 평균과 분산의 범위 내에서 잡음이 추가되므로 원 자료의 유용성을 해치지 않으나, 잡음값은 데이터 값과는 무관하기 때문에 유효한 데이터로 활용하기 곤란하여, 중요한 종적정보는 동일한 잡음을 사용해야함 (예시로 입원일자에 +3이라는 노이즈를 추가하는 경우 퇴원일자에도 +3이라는 노이즈를 부여해야 전체 입원일수에 변화가 없음)

| 생년월일 | 잡음추가 | 잡음추가생년월일 |
|---|---|---|
| 2011-12-05 | +3 | 2011-12-08 |
| 2016-08-09 | -2 | 2016-08-07 |
| 2009-02-11 | -5 | 2009-02-06 |
| 1998-05-27 | -6 | 1998-05-21 |
| 1991-06-18 | +9 | 1991-06-27 |

### ❷ 순열(치환)(Permutation) 수치형데이터 문자형데이터

- 기존 값은 유지하면서 개인이 식별되지 않도록 데이터를 재배열하는 방법
- 개인정보를 다른 행 항목의 정보와 무작위로 순서를 변경하여 전체정보에 대한 변경 없이 특정 정보가 해당 개인과 연결되지 않도록 하는 방법

※ 데이터의 훼손 정도가 매우 큰 기법으로 무작위로 순서를 변경하는 조건 선정에 주의 필요

(설명) 원본과 비교하여 평균 분석 시 전체 재배열은 결과가 다르며
동질집합 내 재배열 결과는 동일

| 지역 | 나이 | 소득금액(원본) | 소득금액(전체 재배열) | 소득금액(동질집합 내 재배열) |
|---|---|---|---|---|
| 서울 | 30대 | 5,987,900 | 2,789,200 | 3,009,600 |
| 서울 | 30대 | 8,169,700 | 4,607,300 | 5,987,900 |
| 서울 | 30대 | 3,009,600 | 5,987,900 | 8,169,700 |
| 나주 | 30대 | 4,607,300 | 2,940,100 | 2,940,100 |
| 나주 | 30대 | 3,560,800 | 8,169,700 | 4,607,300 |
| 나주 | 30대 | 2,940,100 | 5,048,300 | 3,560,800 |
| 세종 | 30대 | 6,088,400 | 3,009,600 | 2,789,200 |
| 세종 | 30대 | 2,789,200 | 3,560,800 | 5,048,300 |
| 세종 | 30대 | 5,048,300 | 6,088,400 | 6,088,400 |

| 원본 분석결과 | 지역 | 서울 | 나주 | 세종 |
|---|---|---|---|---|
| | 평균소득 | 5,722,400 | 3,702,733 | 4,641,967 |

| 전체 재배열 분석결과 | 지역 | 서울 | 나주 | 세종 |
|---|---|---|---|---|
| | 평균소득 | 4,461,467 | 5,048,300 | 4,219,600 |

| 동질집합 내 재배열 분석결과 | 지역 | 서울 | 나주 | 세종 |
|---|---|---|---|---|
| | 평균소득 | 5,722,400 | 3,702,733 | 4,641,967 |

### ❸ 토큰화(Tokenisation) 수치형데이터 문자형데이터

- 개인을 식별할 수 있는 정보를 토큰으로 변환 후 대체함으로써 개인정보를 직접 사용하여 발생하는 개인에 대한 식별 위험을 제거하여 개인정보를 보호하는 기술
- 토큰 생성 시 적용하는 기술은 의사난수생성 기법이나 일방향 암호화, 순서보존 암호화 기법을 주로 사용

| 고객번호 | 이름 | 성별 | 핸드폰번호 | 나이 | 회원등급 | 연간 이용액 |
|---|---|---|---|---|---|---|
| D1304365 | 이공재 | 남 | 010-1234-5678 | 30세 | 2등급 | 3,782,459 |

의사난수 생성기 / 암호화 기법 / 형태보존 암호화

| 고객번호 | 이름 | 성별 | 핸드폰번호 | 나이 | 회원등급 | 연간 이용액 |
|---|---|---|---|---|---|---|
| AD921648 | Wzcd88qdp ekfhandkcosekrn | 남 | 159-6857-6384 | 30세 | 2등급 | 3,782,459 |

### ❹ (의사)난수생성기((P)RNG, (Pseudo) Random Number Generator)

수치형데이터 문자형데이터

- 주어진 입력 값에 대해 예측이 불가능하고 패턴이 없는 값을 생성하는 메커니즘으로 임의의 숫자를 개인정보에 할당

※ 난수는 원칙적으로 규칙적인 배열순서가 없는 임의의 수를 의미하며 컴퓨터는 원천적으로 입력에 의한 처리 결과를 반환하는 것으로 처리의 방법과 입력이 동일하면 항상 동일한 출력이 발생하기 때문에 완전한 난수의 생성은 불가능

### ③ 가명·익명처리를 위한 다양한 기술 (기타 기술)

#### ❶ 표본추출(Sampling) 수치형데이터 문자형데이터

- 데이터 주체별로 전체 모집단이 아닌 표본에 무작위 레코드 추출 등의 기법을 통해 모집단의 일부를 분석하여 전체에 대한 분석을 대신하는 기법
- 확률적 표본추출 방법과 비확률적 표본추출 방법으로 나누어지며, 확률적 표본추출이 통계적 분석에 많이 사용
- 확률적 표본추출: 무작위 표본추출(복원 표본추출, 비 복원 표본추출), 계통적 표본추출, 층화 표본추출, 집락 표본추출 등
- 비확률적 표본부출: 임의 표본추출, 판단 표본추출, 할당 표본추출, 누적 표본추출 등

#### ❷ 해부화(Anatomization) 수치형데이터 문자형데이터

- 기존 하나의 데이터셋(테이블)을 식별성이 있는 정보집합물과 식별성이 없는 정보집합물로 구성된 2개의 데이터셋으로 분리하는 기술

| Record ID | 이름 | 성별 | 나이 | 월 납입금액 | 총 납부금액 |
|---|---|---|---|---|---|
| 1 | 조미선 | F | 33 | 817,250 | 66,300,000 |
| 2 | 홍길병 | M | 61 | 4,559,120 | 327,700,000 |
| 3 | 김영심 | F | 50 | 13,601,564 | 41,300,000 |
| 4 | 이미정 | F | 70 | 979,118 | 64,600,000 |
| 5 | 김경태 | M | 40 | 5,501,809 | 23,549,000 |
| 6 | 유영근 | M | 43 | 609,622 | 13,900,000 |

| Record ID | 이름 | 성별 | 나이 |
|---|---|---|---|
| 1 | 조미선 | F | 33 |
| 2 | 홍길병 | M | 61 |
| 3 | 김영심 | F | 50 |
| 4 | 이미정 | F | 70 |
| 5 | 김경태 | M | 40 |
| 6 | 유영근 | M | 43 |

| Record ID | 월 납입금액 | 총 납부금액 |
|---|---|---|
| 1 | 817,250 | 66,300,000 |
| 2 | 4,559,120 | 327,700,000 |
| 3 | 13,601,564 | 41,300,000 |
| 4 | 979,118 | 64,600,000 |
| 5 | 5,501,809 | 23,549,000 |
| 6 | 609,622 | 13,900,000 |

❸ 재현데이터(Synthetic data) 수치형데이터 문자형데이터

- 원본과 최대한 유사한 통계적 성질을 보이는 가상의 데이터를 생성하기 위해 개인정보의 특성을 분석하여 새로운 데이터를 생성하는 기법

※ 원본 데이터 포함 여부에 따라 완전 재현 데이터(Fully Synthetic Data), 부분 재현 데이터(Partially Synthetic Data), 하이브리드 재현 데이터(Hybrid Synthetic Data)로 구분

❹ 동형비밀분산(Homomorphic secret sharing) 수치형데이터 문자형데이터

- 식별정보 또는 기타 식별가능정보를 메시지 공유 알고리즘에 의해 생성된 두 개 이상의 쉐어(share)*로 대체

* 기밀사항을 재구성 하는 데 사용할 수 있는 하위 집합

※ 재식별은 가명·익명처리된 데이터의 쉐어를 소유한 모두가 동의하는 경우만 가능

❺ 차분 프라이버시(Differential privacy) 수치형데이터 문자형데이터

- 특정 개인에 대한 사전지식이 있는 상태에서 해당정보가 포함된 데이터베이스와 포함되지 않은 데이터베이스 질의(Query)에 대한 응답 값으로 개인을 알 수 없도록 응답 값에 임의의 숫자 잡음(Noise)을 추가하여 특정 개인의 존재 여부를 알 수 없도록 하는 기법
- 1개 항목이 차이나는 두 데이터베이스간의 차이(확률분포)를 기준으로 하는 프라이버시 보호 모델

※ 질의응답 값을 확률적으로 일정 크기 이하의 차이를 갖도록 함으로써 차이에 따른 차분 공격 방지

## 참고2 특이정보 처리 사례

### ① 필요성

- 개인정보를 가명처리를 통해 특정 개인을 알아볼 수 없게 처리했더라도 '특이정보'를 통해 다른 정보와 쉽게 결합하여 개인을 알아 볼 수 있음
  - 따라서, 특이정보의 유형 등을 살펴보고 가명정보 내 해당 유형의 정보가 존재하고 있는지 검토할 필요가 있음

※ 특이정보는 관측된 데이터의 범위에서 많이 벗어난 아주 작은 값이나 아주 큰 값을 의미

### ② 특이정보 사례

- 특정 기관의 급여가 2천만원에서 6천만원까지 고루 분포되어 있는데, 일부 고액 급여 수령자가 발생하는 경우
- 특정 직업의 소속인원이 전국에서 약 300명 정도로 추정되는데, 지역에 극소수(1~2인)만 존재하고 있는 경우
- 정보공개 규정에 따라 공개되는 정보에서 특정 나이대가 현저하게 적게 나타나는 경우

### ③ 특이정보 관찰 방법

- 정보의 특이정보는 3시그마규칙 또는 도수분포표 등을 이용하여 검토할 수 있음
  - 3시그마 규칙 : 68-95-99.7규칙이라고도 하며, 정보의 분포의 3시그마(표준편차) 범위에 거의 모든 값들(99.7%)가 들어가는 것을 의미
  - 도수분포표 : 항목에 대한 값을 적당한 범위로 분류하고, 각 범위에 해당하는 수량을 조사하여 표로 나타내는 것을 의미

• 급여

| 직원 | 급여(만원) |
|---|---|
| 직원1 | 2,200 |
| 직원2 | 3,400 |
| 직원3 | 4,600 |
| 직원4 | 5,300 |
| 직원5 | 10,000 |
| 직원6 | 6,700 |

※ 3시그마 규칙을 이용 하여 표준편차에 벗어난 특이정보 검토

• 지역, 직업

| 주소 | 직업 | 빈도 |
|---|---|---|
| 경기 | 국회의원 | 5 |
| 경기 | 국회의원 | 5 |
| 강원 | 국회의원 | 1 |
| 경기 | 국회의원 | 5 |
| 경기 | 국회의원 | 5 |
| 경기 | 국회의원 | 5 |

※ 지역에 대한 도수분포 (빈도)를 이용하여 특이정보 검토

• 나이

| 나이(세) | 빈도 |
|---|---|
| 10~20 | 4 |
| 20~30 | 11 |
| 30~40 | 21 |
| 40~50 | 18 |
| 50~60 | 5 |
| 60~70 | 1 |

※ 특정 나이에 도수분포(빈도)를 측정 하여 특이정보 검토

### ④ 특이정보 처리 사례

○ 삭제 기법을 활용한 목적별 사용 예시

- 분석 목적에 해당 정보가 없어도 분석에 크게 영향이 없는 경우에만 가능한 기법, 해당 특이정보를 삭제하여 개인 식별성을 제거

**가. 로컬 삭제(Local suppression)**

일반적으로 특이정보 처리에 많이 사용되는 기법으로 도수분포표를 활용하여 빈도가 적은 항목을 삭제하여 처리하는 방법

〈 로컬삭제 기법 예시 〉

| 나이 | 주소 | 직업 | 월소득 |
|---|---|---|---|
| 35 | 서울 | 변호사 | 600만원 |
| 35 | 서울 | 변호사 | 700만원 |
| 35 | 서울 | 변호사 | 500만원 |
| 35 | 서울 | 변호사 | 700만원 |
| 35 | 서울 | 변호사 | 1,200만원 |
| 35 | 경기 | 변호사 | 800만원 |
| 35 | 경기 | 변호사 | 600만원 |
| 35 | 경기 | 변호사 | 1,300만원 |
| 35 | 경기 | 변호사 | 300만원 |
| 35 | 경기 | 변호사 | 900만원 |
| 35 | 경기 | 변호사 | 800만원 |
| 35 | 울릉도 | 변호사 | 200만원 |

| 나이 | 주소 | 직업 | 소득 |
|---|---|---|---|
| 35 | 서울 | 변호사 | 600만원 |
| 35 | 서울 | 변호사 | 700만원 |
| 35 | 서울 | 변호사 | 500만원 |
| 35 | 서울 | 변호사 | 700만원 |
| 35 | 서울 | 변호사 | 1,200만원 |
| 35 | 경기 | 변호사 | 800만원 |
| 35 | 경기 | 변호사 | 600만원 |
| 35 | 경기 | 변호사 | 1,300만원 |
| 35 | 경기 | 변호사 | 300만원 |
| 35 | 경기 | 변호사 | 900만원 |
| 35 | 경기 | 변호사 | 800만원 |
| 35 | Null | 변호사 | 200만원 |

### 나. 행 삭제(Record suppression)

특이정보로 인해 개인의 식별가능성이 있는 경우 사용되는 기법으로 특이정보를 가지고 있는 행 전체를 삭제하여 처리하는 방법

※ 통계 분석에서 특이정보는 분석 목적을 달성하기보다 분석의 목적을 저해하는 요소로 작용하는 경우가 있으며, 이 경우 행 삭제 기법이 가장 적절한 기법이 될 수 있음

〈 레코드 삭제 기법 예시 〉

| 나이 | 주소 | 직업 | 월소득 |
|---|---|---|---|
| 35 | 서울 | 변호사 | 600만원 |
| 35 | 서울 | 변호사 | 700만원 |
| 35 | 서울 | 변호사 | 500만원 |
| 35 | 서울 | 변호사 | 700만원 |
| 35 | 서울 | 변호사 | 1,200만원 |
| 35 | 경기 | 변호사 | 800만원 |
| 35 | 경기 | 변호사 | 600만원 |
| 35 | 경기 | 변호사 | 7,300만원 |
| 35 | 경기 | 변호사 | 300만원 |
| 35 | 경기 | 변호사 | 900만원 |
| 35 | 경기 | 변호사 | 800만원 |
| 35 | 경기 | 변호사 | 200만원 |

| 나이 | 주소 | 직업 | 소득 |
|---|---|---|---|
| 35 | 서울 | 변호사 | 600만원 |
| 35 | 서울 | 변호사 | 700만원 |
| 35 | 서울 | 변호사 | 500만원 |
| 35 | 서울 | 변호사 | 700만원 |
| 35 | 서울 | 변호사 | 1,200만원 |
| 35 | 경기 | 변호사 | 800만원 |
| 35 | 경기 | 변호사 | 600만원 |
| | | | |
| 35 | 경기 | 변호사 | 300만원 |
| 35 | 경기 | 변호사 | 900만원 |
| 35 | 경기 | 변호사 | 800만원 |
| 35 | 경기 | 변호사 | 200만원 |

○ 통계적 기법의 종류와 목적별 사용 예시

- 분석 목적에 특이정보를 가지고 있는 해당 정보가 필요한 경우 활용하는 기법으로, 해당 특이정보를 통계적인 방법을 통해 통계값으로 변경하여 사용

**가. 단일 속성으로 대체(Combining a set of attributes into a single attribute)**

숫자형 정보가 아닌 경우(문자형 등) 주로 사용되는 방법으로 분류군의 상위로 묶어 처리하는 방법

※ 특정한 직업이 희귀하여 개인의 식별이 가능한 경우 상위의 분류로 변경하여 사용함으로 희귀성을 제거

〈 단일속성 대체 예시 〉

| 나이 | 주소 | 직업 | 월소득 |
|---|---|---|---|
| 35 | 서울 | 변호사 | 600만원 |
| 35 | 서울 | 변호사 | 700만원 |
| 35 | 서울 | 변호사 | 500만원 |
| 35 | 서울 | 변호사 | 700만원 |
| 35 | 서울 | 판사 | 1,200만원 |
| 35 | 경기 | 검사 | 800만원 |
| 35 | 경기 | 변호사 | 600만원 |
| 35 | 경기 | 변호사 | 1,300만원 |
| 35 | 경기 | 변호사 | 300만원 |
| 35 | 경기 | 변호사 | 900만원 |
| 35 | 경기 | 변호사 | 800만원 |
| 35 | 경기 | 변호사 | 200만원 |

| 나이 | 주소 | 직업 | 소득 |
|---|---|---|---|
| 35 | 서울 | 변호사 | 600만원 |
| 35 | 서울 | 변호사 | 700만원 |
| 35 | 서울 | 변호사 | 500만원 |
| 35 | 서울 | 변호사 | 700만원 |
| 35 | 서울 | 법조인 | 1,200만원 |
| 35 | 경기 | 법조인 | 800만원 |
| 35 | 경기 | 변호사 | 600만원 |
| 35 | 경기 | 변호사 | 1,300만원 |
| 35 | 경기 | 변호사 | 300만원 |
| 35 | 경기 | 변호사 | 900만원 |
| 35 | 경기 | 변호사 | 800만원 |
| 35 | 경기 | 변호사 | 200만원 |

### 나. 로컬 일반화(Local generalization)

선택한 행에서 일부 특정 값을 일반화하여 활용하는 기법으로, 다른 행의 속성값은 수정하지 않고 희귀 값을 가진 속성값만 처리하여 사용

〈 로컬 일반화(상단 코딩) 기법 예시 〉

| 나이 | 주소 | 직업 | 월소득 |
|---|---|---|---|
| 35 | 서울 | 변호사 | 600만원 |
| 35 | 서울 | 변호사 | 700만원 |
| 35 | 서울 | 변호사 | 500만원 |
| 35 | 서울 | 변호사 | 700만원 |
| 36 | 서울 | 변호사 | 1,200만원 |
| 36 | 경기 | 변호사 | 800만원 |
| 36 | 경기 | 변호사 | 600만원 |
| 36 | 경기 | 변호사 | 1,300만원 |
| 37 | 경기 | 변호사 | 300만원 |
| … | … | … | … |
| 84 | 경기 | 변호사 | 800만원 |
| 88 | 경기 | 변호사 | 200만원 |

| 나이 | 주소 | 직업 | 소득 |
|---|---|---|---|
| 35 | 서울 | 변호사 | 600만원 |
| 35 | 서울 | 변호사 | 700만원 |
| 35 | 서울 | 변호사 | 500만원 |
| 35 | 서울 | 변호사 | 700만원 |
| 36 | 서울 | 변호사 | 1,200만원 |
| 36 | 경기 | 변호사 | 800만원 |
| 36 | 경기 | 변호사 | 600만원 |
| 36 | 경기 | 변호사 | 1,300만원 |
| 37 | 경기 | 변호사 | 300만원 |
| … | … | … | … |
| 80초과 | 경기 | 변호사 | 800만원 |
| 80초과 | 경기 | 변호사 | 200만원 |

### 다. 부분 총계(Micro Aggregation)

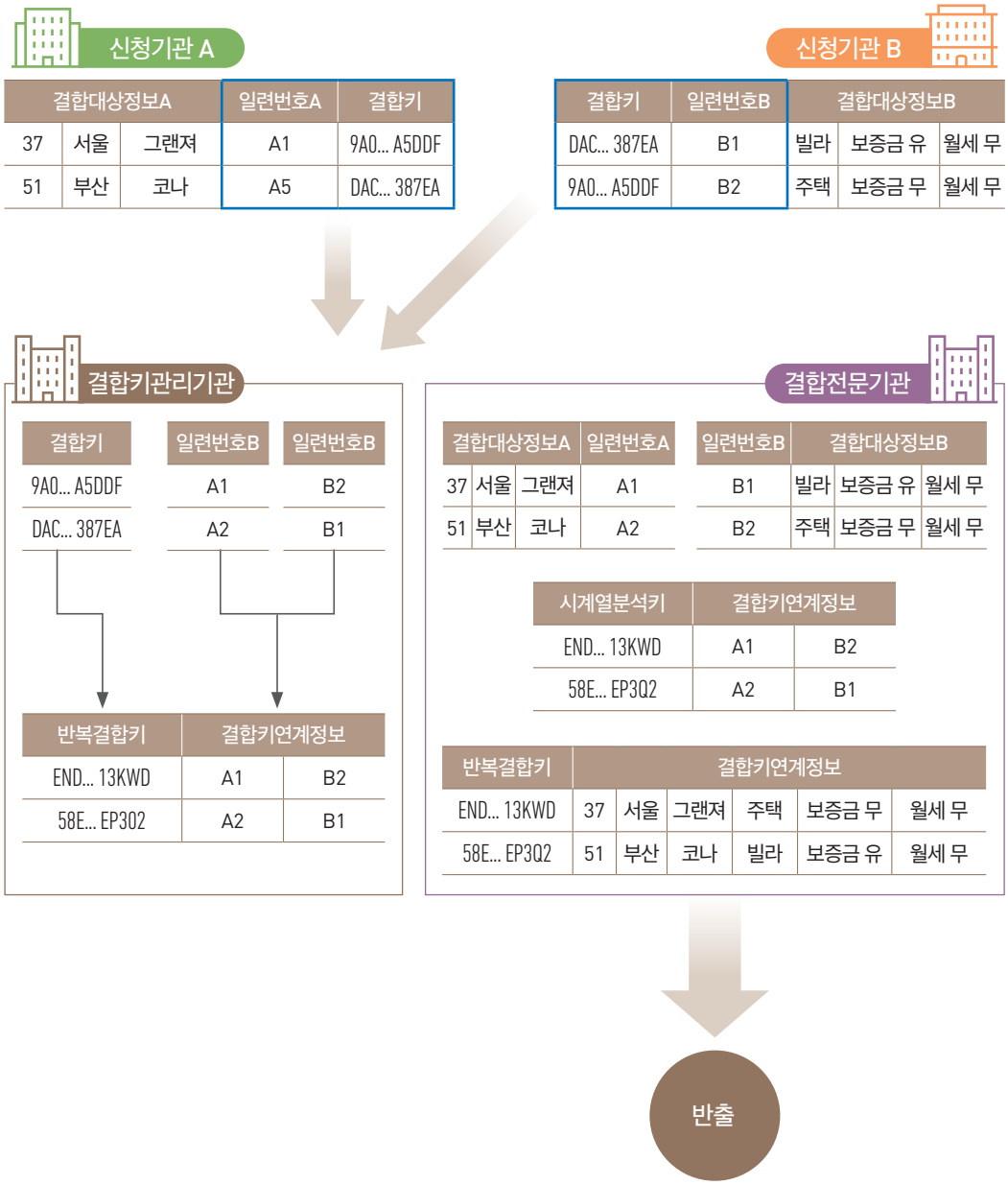
부분 총계는 일부(특정그룹 값의 합)속성에서 정확한 통계적 값을 확인하는 기법으로,
로컬일반화 보다 일부 속성에서 정확한 값을 알 수 있음

〈 부분 총계 기법 예시 〉

| 나이 | 주소 | 직업 | 월소득 |
|---|---|---|---|
| 35 | 경기 | 변호사 | 600만원 |
| 35 | 경기 | 변호사 | 700만원 |
| 35 | 경기 | 변호사 | 500만원 |
| 35 | 경기 | 변호사 | 700만원 |
| 35 | 경기 | 변호사 | 6,200만원 |
| 35 | 경기 | 변호사 | 800만원 |
| 35 | 경기 | 변호사 | 600만원 |
| 35 | 경기 | 변호사 | 7,300만원 |
| 35 | 경기 | 변호사 | 300만원 |
| 35 | 경기 | 변호사 | 900만원 |
| 35 | 경기 | 변호사 | 800만원 |
| 35 | 경기 | 변호사 | 200만원 |

| 나이 | 주소 | 직업 | 소득 |
|---|---|---|---|
| 35 | 경기 | 변호사 | 600만원 |
| 35 | 경기 | 변호사 | 700만원 |
| 35 | 경기 | 변호사 | 500만원 |
| 35 | 경기 | 변호사 | 700만원 |
| 35 | 경기 | 변호사 | 6,750만원 |
| 35 | 경기 | 변호사 | 800만원 |
| 35 | 경기 | 변호사 | 600만원 |
| 35 | 경기 | 변호사 | 6,750만원 |
| 35 | 경기 | 변호사 | 300만원 |
| 35 | 경기 | 변호사 | 900만원 |
| 35 | 경기 | 변호사 | 800만원 |
| 35 | 경기 | 변호사 | 200만원 |

## 참고3 시계열 분석을 위한 반복결합 절차



① 각 결합신청자는 결합대상정보의 정보주체에 대한 일련번호를 생성하고, 결합키관리기관과 합의하여 정해진 방법에 따라 결합키를 생성

- 생성된 결합키와 일련번호는 결합키관리기관으로 송신

② 결합키관리기관은 각 결합신청자로부터 수신받은 결합키를 활용하여 1) 시계열 분석에 필요한 키(반복결합키)와 2) 결합키연계정보를 생성

※ 일반적인 결합의 경우, 추가적인 시계열 분석에 필요한 키를 생성하지 않음

③ 결합키관리기관에서 생성한 1), 2)의 정보를 결합전문기관에 송신하고, 결합전문기관은 수신받은 2)를 활용하여 각 결합신청자의 가명정보를 결합

④ 결합전문기관은 결합된 가명정보에 1)을 포함하여 반출하고, 결합신청자는 반출된 정보에 대한 안전조치 의무 수행

※ 유의사항 : 반복결합 신청자는 추후 반출되는 정보와의 연계·분석을 위하여 결합키에 사용된 결합키 생성항목, 인코딩 방식, 알고리즘(Salt값 제외)을 보관하여야 함

**〈 추가 시계열 분석 신청 및 활용 방법 〉**

결합신청자가 추가 시계열 분석을 위한 반복결합을 신청하는 경우, 결합신청자는 최초 반복결합 신청시 사용한 방식에 따라 결합키를 생성하여 일련번호와 함께 결합키관리기관에 전달하고, 결합 후 반출된 반출정보에 포함된 1)의 정보를 활용하여 내부에서 연계하여 활용

※ 시계열 분석이 완전히 종료된 경우 이를 결합키관리기관에 통하여야 하며, 결합키관리기관은 해당 결합 후 보관하고 있는 1)의 생성방법을 삭제

## 참고4 가명처리 및 결합 목적 증빙 자료 예시

### 1. 통계작성 계획서 예시

**통계작성 계획서**

| | | |
|---|---|---|
| 통계명 | | |
| 대표 참여진 | 소속 | |
| | 담당자명 | |
| 통계작성 배경 및 목적 | | |
| 통계작성 대상자 수 | | |
| 통계작성 계획 및 방법 | | |
| 기대효과 및 활용방안 | | |
| 붙임. 상세 통계작성 계획서 등 | | |

## 2. 과학적 연구 계획서 예시

| 과학적 연구 계획서 | | |
|---|---|---|
| 연구명 | | |
| 연구진 | 소속 | |
| | 연구책임자 | |
| 연구 배경 및 목적 | | |
| 예상 연구 기간 | | |
| 연구 대상자 수 | | |
| 연구 방법 | | |
| 연구내용 | | |
| 기대효과 및 활용방안 | | |
| 붙임. 상세 연구계획서 등 | | |

### 3. 공익적 기록보존 계획서 예시

| 공익적 기록보존 계획서 | | |
|---|---|---|
| 공익적 기록보존명 | | |
| 대표 참여진<br>(기록보관 기관) | 보관기관명 | |
| | 담당자명 | |
| 공익적<br>기록보존 목적 | | |
| 보존기간 | | |
| 공익적<br>기록보존 방법 | | |
| 내용 | | |
| 기대효과 및<br>활용방안 | | |
| 붙임. 상세 계획서 등 | | |

## 참고5 결합신청서 작성 방법

### ① 결합신청서

② 신청번호 / 접수번호

☐ 가명정보 제공 ☐ 가명정보 제공+결합정보 이용 ☐ 결합정보 이용

| 결합신청자 | | | |
|---|---|---|---|
| 기관명 | A사 | 사업자등록번호 또는 법인등록번호 | ○○○-○○-○○○○○ |
| 주소 | ○○시 ○○구 ○○○ | 대표자명 | ○○○ |
| 담당자 | 홍길동 | 담당자 연락처 (전화, e-mail) | 010-○○○○-○○○○ ○○○○○@○○○○○.○○.○○ |
| 유형 | [ ] 개인 [ ∨ ] 공공기관 [ ] 비영리법인 [ ] 민간기관 | | |

| 결합 개요 | |
|---|---|
| ③ 반복결합 | [ ∨ ] 해당없음 [ ] 최초 [ ] 추가(결합접수번호: ) |
| ④ 추가절차 신청 | 결합률 확인 [ ∨ ] 가명정보 추출 [ ∨ ] 모의결합 [ ] |

| ⑤ 가명정보 제공자 | 해당없음 [ ] |
|---|---|
| 파일명 | abc |
| 제출 방법 | [ ∨ ] 온라인 [ ] 오프라인 |
| 제출 예정일 | ○○○○년 ○○월 ○○일 |
| 제공정보 요약 | 파일 크기(○.○GB) 전체 레코드 수(○○○,○○○개) 모의결합 레코드 수(○○○개) |
| 전체 가명정보 제공자명(총수) | 총 2개: A사(파일명: abc), B사(파일명: zyx) |
| 지원 요청 사항 | [ ∨ ] 결합 신청에 필요한 가명처리 |

| ⑥ 결합정보 이용자 | 해당없음 [ ] |
|---|---|
| 결합 목적 | [ ] 통계작성 [ ∨ ] 과학적 연구 [ ] 공익적 기록보존 등 |
| 세부 결합 목적 | 구체적 목적 설명 |
| 분석공간 이용 | [ ] 추가 가명처리만 수행 [ ∨ ] 결합정보 분석 [ ] 이용안함 |
| 지원 요청 사항 | [ ∨ ] 반출 전 처리 [ ∨ ] 분석 |

「개인정보 보호법」 제28조의3제1항 및 같은 법 시행령 제29조의3제1항에 따른 결합을 위하여 결합전문기관에 결합신청서를 위와 같이 제출합니다.

○○○○년 ○○월 ○○일

결합신청자 홍길동 (서명 또는 인)

**결합전문기관의 장** 귀하

⑦ 첨부 서류

1. 사업자등록증, 법인등기부등본 등 결합신청자 관련 서류 1부
2. 결합 목적을 증명할 수 있는 서류 1부(결합된 정보를 반출하려는 자에 한함)
3. 결합 대상 가명정보에 관한 서류(전체 항목명, 가명처리 대상 항목명*, 가명처리 내역 등**) 1부(가명정보 제공자에 한함)
   * 결합키 생성에 사용된 항목 제외
   ** 결합대상정보가 확정된 이후에 제출

### ① 결합신청서 작성

- (작성 주체) 결합신청서는 결합신청자*별로 각자 제출하는 것이 원칙

  * 가명정보를 보유하고 있는 개인정보처리자, 현재 가명정보를 보유하고 있지 않으나 결합된 가명정보를 처리할 예정인 개인정보처리자

  - 작성자가 해당되는 결합신청자의 유형(가명정보 제공, 가명정보 제공 및 결합정보 이용, 결합정보 이용) 표기

### ② 신청번호 및 접수번호

- (신청번호) 결합신청자 중 대표자*가 결합종합지원시스템(link.privacy.go.kr)을 통해 발급받은 번호

  * 결합 신청에 있어 총괄 관리·감독을 수행할 결합신청자

- (접수번호) 결 제출된 결합신청서를 접수할 때 발행하는 번호, 결합신청서 제출 시 공란으로 제출

### ③ 반복결합

- 추후 동일한 목적/형태 등으로 주기적·반복적 결합을 수행하는 경우

  - (최초) 반복결합을 최초로 신청하는 경우

  - (추가) 최초 반복결합이 완료된 이후 반복결합을 추가로 신청한 경우, 최초 반복결합 신청시 발급되었던 결합 접수번호 기재

### ④ 추가절차 신청

- 결합률 확인, 추출, 모의결합*을 신청하는 자는 해당 사항을 체크(중복체크 가능, 선택사항)

  * 모의결합의 경우 결합전문기관별로 지원여부가 다르므로 신청하려는 결합전문기관이 모의결합을 지원하는지 여부를 확인한 후 신청 필요

### ⑤ 가명정보 제공자(개인정보처리자)

- 결합을 위해 가명정보를 결합전문기관에 제공하는 자가 작성하며, 가명정보를 보유하고 있지 않은 자는 해당없음에 체크하고 나머지 항목은 공란

  - 가명정보 파일명, 제출 방법, 제공 정보 요약, 전체 가명정보 제공자명(총수)*, 지원 요청 사항** 등 작성

  * 해당 결합을 신청하는 가명정보 제공자의 전체 기관명 및 전체 기관수(총 ○개)

  ** 가명처리를 직접 수행하기 어려운 가명정보 제공자는 결합전문기관에 결합 전 가명처리의 지원을 요청할 수 있음(이 경우 해당 결합전문기관의 지원여부 확인 필요)

### ⑥ 결합정보 이용자(개인정보처리자)

- 결합된 정보를 이용하려는 자(현재 가명정보를 보유하고 있지 않은 자 포함)가 작성하며, 가명정보를 제공하나 결합된 정보를 이용하지 않는 자는 해당없음에 체크하고 공란
  - 결합 목적, 분석공간 이용여부*, 지원 요청 사항(중복체크 가능)** 작성

    * 결합된 정보를 결합전문기관이 제공하는 인프라를 활용하여 추가처리 및 분석을 하고자 하는 경우 (선택사항) 체크

    ** 결합전문기관에 결합된 정보에 대한 반출 전 가명·익명처리, 결합된 정보를 분석을 요청하고자 하는 자는 해당 지원 사항란 표시(이 경우 해당 결합전문기관의 지원여부 확인 필요)

### ⑦ 첨부 서류

- 각 첨부 서류별 제출 주체는 아래와 같으며, 가명정보를 제공하는 자가 결합된 정보도 처리하고자 하는 경우에는 모든 첨부 서류 제출 필요
  - 첨부 서류는 결합신청 시 제출하는 것이 원칙이나 결합전문기관과 협의를 통해 결합률 확인, 가명정보 추출 등 모든 사전절차 완료되어 결합 대상이 확정된 이후 결합 대상 가명정보에 관한 서류를 제출할 수 있도록 협의 가능

| | 가명정보 제공자 | 결합정보 이용자 |
|---|---|---|
| 1. 결합신청자 관련 서류 | ○ | ○ |
| 2. 결합 목적 관련 서류 | | ○ |
| 3. 결합 대상 가명정보에 관한 서류 | ○ | |

## 참고6 반출신청서 작성 방법

# ① 반출신청서

② 

| 반출접수번호 | |
|---|---|
| 결합접수번호 | |

| 결합신청자 | | | |
|---|---|---|---|
| 기관명 | A사 | 사업자등록번호<br>또는 법인등록번호 | ○○○-○○-○○○○○ |
| 주소 | ○○시 ○○구 ○○○ | 대표자명 | ○○○ |
| 담당자 | 홍길동 | 담당자 연락처<br>(전화, e-mail) | 010-○○○○-○○○○<br>○○○○○@○○○○○.○○.○○ |

| ③ 결합 유형 | |
|---|---|
| 반복결합 | [ ] 최초 [ ] 추가 |

| ④ 반출 개요 | |
|---|---|
| 파일명 | ccdab |
| 반출 목적 | [ ] 통계작성 [ ∨ ] 과학적 연구 [ ] 공익적 기록보존 등 |
| 세부 반출 목적 | 구체적 목적 설명 |
| 반출정보 유형 | [ ∨ ] 가명정보 [ ] 법 제58조의2에 해당하는 정보(익명정보) |
| 제공 받는 방법 | [ ∨ ] 온라인 [ ] 오프라인 [ ] 결합전문기관 내 분석공간 |
| 지원 요청 사항 | [ ∨ ] 반출된 정보의 분석 [ ∨ ] 개인정보 보호 교육 |

「개인정보 보호법」 제28조의3제2항 및 같은 법 시행령 제29조의3제3항·제6항,
「가명정보의 결합 및 반출 등에 관한 고시」 제10조제3항에 따라 결합된 정보를 반출하기 위하여 결합전문기관에 반출신청서를 위와 같이 제출합니다.

○○○○년 ○○월 ○○일

결합신청자 홍길동 (서명 또는 인)

**결합전문기관의 장** 귀하

| ⑤ 첨부 서류 | 1. 반출 대상 정보에 관한 서류 1부(추가적인 서류 제출이 필요한 경우에 한함)<br>2. 반출 목적을 증명할 수 있는 서류 1부(추가적인 서류 제출이 필요한 경우에 한함)<br>3. 반출정보의 안전조치계획 및 이를 증빙할 수 있는 서류 1부 |
|---|---|

### ① 반출신청서 작성

- (작성 주체) 결합된 정보 또는 분석결과 등을 결합전문기관 외부로 반출하고자하는 자는 반출신청서를 작성하여 결합전문기관에 제출

  ※ 가명정보를 제공만 하는 자와 결합된 정보를 결합전문기관 내의 분석공간에서 분석만을 수행하는 자는 반출신청서를 작성하지 않아도 됨

### ② 반출접수번호 및 결합접수번호

- (반출접수번호) 결합전문기관이 제출된 반출신청서를 접수할 때 발행하는 번호, 반출신청서 제출 시 공란으로 제출
- (결합접수번호) 결합전문기관이 제출된 결합신청서를 접수할 때 발행하였던 번호

### ③ 결합 유형

- 반복결합을 신청한 자는 최초/추가 여부를 체크하며, 반복결합이 아닌 경우 공란

### ④ 반출 개요

- 결합된 정보를 반출하려는 결합신청자는 파일명(반출할 결합 결과물), 반출 목적, 반출정보 유형 등을 작성

  ※ 결합신청자는 반출심사 전인 경우 결합 목적과 반출 목적 변경 가능

  - (반출정보 유형) 반출을 신청하려는 자가 정보의 형태 등을 고려하여 가명 또는 익명으로 판단하여 표기
  - (제공받는 방법) 결합전문기관등이 제공하는 시스템을 통해 제공받는 경우는 온라인, USB등 저장장치를 이용해 반출하는 경우는 오프라인, 결합전문기관이 제공하는 분석공간을 이용하는 경우로 구분하여 표기
  - (지원 요청사항) 정보의 분석을 위해 결합전문기관의 지원이나 가명정보의 처리에 관한 교육이 필요하면 표기

### ⑤ 첨부 서류

- 추가적인 서류 제출이 필요한 경우에 한하여 모든 서류를 제출
  - (반출 대상 정보에 관한 서류) 분석공간을 통해 추가 가명처리가 수행되어 반출 대상 정보가 당초 제출한 결합 대상 가명정보와 상이한 경우에만 제출하고 동일한 경우에는 생략 가능
  - (반출 목적 관련 서류) 반출 목적이 당초의 결합 목적과 달라진 경우(결합 목적과 반출 목적의 양립 가능성 검토 필요)만 제출하고 동일한 경우에는 생략 가능
  - (안전조치 계획) 개인정보 처리방침, 내부 관리계획, 운영 지침 등 반출정보의 안전조치와 관련된 자료를 제출

## 참고7 내부 관리계획 작성 예시

**제00조(가명정보 및 추가정보 관리책임자 지정)** ① ○○○○○(개인정보처리자명)는 가명정보에 대한 총괄 관리책임자로 ○○○○○(가명정보 관리책임자명 또는 직책)로 정한다.

② 가명정보 관리책임자는 다음과 같은 역할을 수행한다.

1. 가명정보에 대한 내부 관리계획의 수립·시행
2. 내부 관리계획의 이행실태 점검 및 관리
3. 가명처리 및 적정성 검토 현황 관리
4. 가명정보 및 추가정보에 대한 관리·감독
5. 가명정보 처리 현황 및 관련 기록 관리
6. 가명정보를 처리하는 자 교육계획의 수립 및 시행
7. 가명처리 및 가명정보 처리 위탁 사항에 대한 관리·감독(해당 시)
8. 가명정보에 대한 재식별 모니터링 및 재식별 시 처리 방안의 수립·시행
9. 그 밖의 가명정보 처리에 대한 보호에 관한 사항

**제00조(가명정보 및 추가정보의 분리보관)** ① 가명정보는 가명처리가 완료되면 가명처리 전 개인정보와 분리·보관하여야 한다.

② 가명처리의 과정에서 발생하는 추가정보는 가명정보와 분리·보관하여야 한다.

③ 가명처리 전 개인정보, 가명정보 및 추가정보는 물리적으로 분리 보관하는 것을 원칙으로 하며 물리적 보관이 어려운 경우 논리적인 분리를 시행할 수 있다.

④ 논리적으로 분리·보관하는 경우 엄격한 접근통제를 적용해야 한다.

**제00조(가명정보 및 추가정보에 대한 접근권한 분리)** ① 가명처리가 완료되면 가명정보 또는 추가정보의 접근권한은 최소한의 인원으로 엄격하게 통제하여야 하며, 업무에 따라 차등적으로 부여 하여야 한다.

② 추가정보에 대한 접근권한과 가명정보에 대한 접근권한은 분리하여 관리해야 한다.
③ 가명정보 또는 추가정보에 대한 접근권한 부여, 변경 또는 말소에 대한 내역을 기록하도록 하고 이 기록은 최소 3년간 보관하여야 한다.

**제00조(가명정보 및 추가정보의 안전성 확보조치)** ① 가명정보와 추가정보는 개인정보보호법 및 동법 시행령에서 요구하는 안전성 확보조치를 수행하여야 한다.
② 추가정보에 특별한 이유가 없는 한 생성 즉시 삭제하도록 한다. 단, 시계열 분석 등의 이유로 추가정보가 필요한 경우 저장 시 암호화하여 저장하여야 한다.

**제00조(가명정보를 처리하는 자의 교육)** ① 가명정보 관리책임자는 가명정보를 처리하는 자에게 필요한 가명정보 보호 교육계획을 수립하고 실시하여야 한다.
② 가명정보 보호 교육은 다음과 같은 내용을 포함하여 시행하여야 한다.
1. 가명정보 처리 근거에 관한 사항
2. 가명정보 및 추가정보의 안전조치에 관한 사항
3. 재식별 금지에 관한 사항
③ 가명정보를 처리하는 자에 대한 교육은 개인정보 보호교육과 함께 수행할 수 있으며 교육을 실시한 결과 또는 이를 입증할 수 있는 관련 자료 등을 기록·보관하여야 한다.

**제00조(가명정보 처리 기록 작성 및 보관)** ① 가명정보의 처리 시 다음과 같은 사항에 대해 가명정보 처리 대장에 기록을 작성하여 보관하여야 한다.
1. 가명정보의 처리 목적
2. 가명처리한 개인정보의 항목
3. 가명정보의 이용내역
4. 제3자 제공 시 제공받는 자
5. 그 밖에 가명정보의 처리 내용을 관리하기 위하여 개인정보보호위원회가 필요하다고 인정하여 고시하는 사항

**제00조(개인정보 처리방침 공개)** ① 가명정보 처리와 관련하여 아래와 같은 내용을 개인정보 처리방침에 포함하여 공개하여야 한다.

1. 가명정보의 처리 목적
2. 가명정보 처리기간(선택)
3. 가명정보 제3자 제공에 관한 사항(해당 시)
4. 가명정보 처리 위탁에 관한 사항(해당 시)
5. 처리하는 가명정보의 항목
6. 가명정보의 안전성 확보조치에 관한 사항

**제00조(가명정보의 재식별 금지)** ① 가명정보를 처리하는 자의 가명정보에 대한 재식별 행위는 엄격하게 금지한다.

② 가명정보를 처리하는 자가 가명정보를 처리하는 중 특정 개인에 대한 재식별이 발생하는 경우 즉시 처리를 중단하고 이를 가명정보 관리책임자에게 통보한 후 수립된 재식별 시 처리 방안에 따라 즉시 조치하여야 한다.

# 가명정보 처리 가이드라인

2021년 10월 발행
발행처: 개인정보보호위원회
지원기관: 한국인터넷진흥원



• 본 가이드는 개인정보보호포털(https://privacy.go.kr)에서
무료로 다운받으실 수 있습니다.

본 가이드라인은 2021년 10월 기준으로 작성되었습니다.
항상 최신의 가이드라인은 개인정보보호위원회 개인정보보호
포털(자료마당→지침자료)에서 확인하시기 바랍니다.


개인정보보호위원회
Personal Information Protection Commission