개인정보 비식별 가이드 라인

- 비식별 조치 기준 및 지원·관리체계 -

NIA 한국정보화진흥원

목 차

▣ 교육 개요

빅데이터 활용에 필요한 비식별 조치 기준·절차·방법 등을 구체적으로 안내하여 안전한 빅데이터 활용기반 마련과 개인정보 보호 강화를 도모

▣ 목 차

1. 개인정보 보호 관련 법령 통합 해설서

1. 개인정보 보호 관련 법령 통합 해설서

▣ 개 요 (1)

○ 개인정보의 범위(법 제2조 1호 관련) 명확화

개인정보 범위를 **'개인을 알아볼 수 있는 정보'**와 **'다른 정보와 쉽게 결합하여'** 알아볼 수 있는 정보

·'알아볼 수 있는'의 주체를 **'해당정보를 처리하는 자'**로 명확화
·다른 정보와 결합 시 '단순 쉽게'가 아닌 **정보의 입수가능성, 결합가능성도 높아야**함을 명시

○ 비식별 조치에 대한 근거 마련

개인정보 비식별 조치의 **법적 의미와 효과에 대한 명확한 해석 부재**

· 개인 **식별요소 제거조치 시 비식별 정보**로 규정
· **비식별 조치한 정보는 개인정보가 아닌 것**으로 추정
· 재식별 가능성을 방지하고자 **필수 보호조치 준수** 必

1. 개인정보 보호 관련 법령 통합 해설서

▣ 개 요 (2)

○ 현행법상 제재수단 안내

· 비식별 정보를 재식별하여 이용하거나 제3자 제공의 경우 [개인정보 목적 외 이용]

☞ **5년 이하 징역 또는 5천만원 이하 벌금**

· 재식별된 개인정보를 파기하지 않고 보관한 경우 [정보주체의 미동의 개인정보 수집]

☞ **5천만원 이하의 과태료**

1. 개인정보 보호 관련 법령 통합 해설서

▣ 개인정보 개념

- 「개인정보 보호법」과 「정보통신망 이용 촉진 및 정보보호 등에 관한 법률(이하 정보통신망법)」에서는 개인정보의 개념을 규정
- 「신용정보의 이용 및 보호에 관한 법률(이하 신용정보법)」에서는 개인신용정보와 개인식별정보의 개념에 대해 규정

✓ 개인정보 보호법과 정보통신망법에서의 개인정보 개념정의는 법률상 표현이 조금 다르게 되어 있으나, 법률 해석상 그 내용은 사실상 동일

✓ 신용정보법 상의 개인신용정보 및 개인식별정보는 개인정보 보호법과 정보통신망법에서 말하는 개인정보 개념과 다르지 않음

1. 개인정보 보호 관련 법령 통합 해설서

▣ 개인정보의 구체적 판단 기준

- ‘생존하는’ 개인에 관한 정보이여야 함
- ‘개인에 관한’ 정보이여야 함
- ‘정보’의 내용·형태 등은 제한이 없음
- 개인을 ‘알아볼 수 있는’ 정보이여야 함
- 다른 정보와 ‘쉽게 결합하여’ 개인을 알아볼 수 있는 정보도 포함

1. 개인정보 보호 관련 법령 통합 해설서

▣ 개인정보 개념 관련 판례 및 유권해석 사례

판례

☞ **휴대전화번호 뒤 4자리**

대전지법 논산지원(2013고단17 판결)은 「휴대전화번호 뒷자리 4자」에 대하여, "휴대전화번호 뒷자리 4자만으로도 그 전화번호 사용자가 누구인지를 식별할 수 있는 경우가 있고, 특히 그 전화번호 사용자와 일정한 인적 관계를 맺어온 사람이라면 더더욱 그러할 가능성이 높으며, 설령 휴대전화번호 뒷자리 4자만으로는 그 전화번호 사용자를 식별하지 못한다 하더라도 그 뒷자리 번호 4자와 관련성이 있는 다른 정보(생일, 기념일, 집 전화번호, 가족 전화번호, 기존 통화내역 등)와 쉽게 결합하여 그 전화번호 사용자가 누구인지를 알아볼 수도 있다"고 하여 「개인정보보호법」 제2조제1호에 규정된 개인정보에 해당된다고 판시하고 있다

유권해석 사례

☞ **개인정보보호 위원회 결정 :** 배달음식점 고객의 전화번호 및 주소

개인정보보호위원회는 2012년 1월 30일 「개인정보보호법 관련 법령해석 요청 건(의안 제2호)」에 대한 의결 이유에서 "배달음식점 고객의 전화번호 및 주소는 그 자체로는 특정 개인을 식별할 수 없지만, 용이하게 다른 정보와 결합하여 특정 개인을 식별할 수 있으므로, 「개인정보보호법」 제2조 제1호의 '개인정보'에 해당함"이라고 해석하고 있다. 용이하게 다른 정보와 결합하여 특정 개인을 식별할 수 있기만 하면 그 자체로서 특정 개인을 식별할 수 없는 경우에도 개인정보로 보고 있으므로, 의결 이유에서 지적한 '고객의 전화번호 및 주소'이외에도 개인정보로 인정할 수 있는 정보의 범위가 확장될 수 있다고 해석할 수 있다.

1. 개인정보 보호 관련 법령 통합 해설서

▣ 비식별 정보의 개념

1. **(비식별 정보의 개념)** 개인정보를 비식별 조치한 정보, 즉 '비식별 정보'란 정보의 집합물에 대해 「개인정보 비식별 조치 가이드라인」에 따라 적정하게 '비식별 조치'된 정보

   - '비식별 조치'란 정보의 집합물에서 개인을 식별할 수 있는 요소를 전부 또는 일부 삭제하거나 대체 등의 방법을 통해 개인을 알아볼 수 없도록 하는 조치

2. **(비식별 정보의 활용)** 비식별 정보는 개인정보가 아닌 정보로 추정되므로 정보주체로부터의 별도 동의없이 해당 정보를 이용하거나 제3자에게 제공할 수 있음

3. **(비식별 정보의 보호)** 비식별 정보는 개인정보가 아닌 것으로 추정되지만, 새로운 결합기술이 나타나거나 결합 가능한 정보가 증가하는 경우에는 정보주체가 '재식별'될 가능성이 있음. 따라서 비식별 정보라고 하더라도 필수적인 관리적·기술적 보호조치는 이행해야 함

1. 개인정보 보호 관련 법령 통합 해설서

▣ 재식별 시 법적 제재

○ 형사처벌

- 비식별 정보를 재식별하여 이용하거나 제3자에게 제공한 경우

- 개인정보의 목적 외 이용·제공에 해당(개인정보 보호법 제18조제1항 위반, 정보통신망법 제24조 및 제24조의2 위반, 신용정보법 제32조 및 제33조 위반)
- 5년 이하의 징역 또는 5천만원 이하의 벌금
  ※ 정보통신망법 적용 사업자는 위반행위 관련 매출액의 3% 이내 과징금 추가 부과

○ 행정처분

- 비식별 정보를 활용하여 재식별하고도 즉시 파기 조치하지 않고 보관하고 있는 경우

- 정보주체의 동의없이 개인정보를 수집한 경우에 해당(개인정보 보호법 제15조제1항 위반, 정보통신망법 제22조제1항 위반, 신용정보법 제15조제2항 위반)
- 5천만원 이하의 과태료가 부과
  ※ 정보통신망법 적용 사업자는 5년 이하 징역 또는 5천만원 이하 벌금형에 처해질 수 있으며 위반행위 관련 매출액의 3% 이내 과징금 추가 부과

2. 개인정보 비식별 조치 가이드라인

2. 개인정보 비식별 조치 가이드라인

▣ 가이드라인 추진 배경

1 정부 3.0 및 빅데이터 활용 확산에 따른 데이터 활용가치 증대

- 공공정보 개방 · 공유는 투명하고 효율적인 정부 운영에, 빅데이터 활용은 과학적 정책 집행 및 맞춤형 서비스 제공에 필수적인 수단
- 특히, 빅데이터 분석, IoT 기술 등을 통한 새로운 서비스 창출과 신산업 활성화에 데이터의 활용가치 증대

2 개인정보 보호 강화에 대한 사회적 요구 지속

- 크고 작은 개인정보 유출 사고가 지속되어 개인정보 보호 정책을 강화해야 한다는 사회적 요구가 계속
- 다양한 데이터 활용을 필요로 하는 새로운 산업과 기술 발전으로 개인정보 침해 위험도 증가 추세

3 '보호와 활용'을 동시에 모색하는 세계적 정책변화에 적극 대응

- 미국 · 영국 등 주요 선진국은 개인정보 침해가능성을 최소화하면서 데이터 산업 활성화를 위한 정책 추진 중
- 사생활 침해 방지를 위한 안전장치 마련과 동시에 비식별 조치된 정보는 산업적으로 활용할 수 있도록 구체적인 가이드 제시 필요

2. 개인정보 비식별 조치 가이드라인

▣ 비식별화 개념

○ 정의

- 데이터 내에 개인을 식별할 수 있는 정보가 있는 경우, 이의 일부 또는 전부를 삭제, 또는 일부를 속성 정보로 대체 처리함으로써 다른 정보와 결합하여도 특정 개인을 식별하기 어렵도록 하는 조치

<비식별화 처리 예시>

▶ 정보 내 식별 가능한 특징을 제거하거나 변형시킴으로써 데이터 집합과 데이터 대상(정보이용자)과의 유일한 연관관계를 제거
  - 개인과 여러 정보를 연결시켜 개인의 정보가 드러나지 않게 하거나 하나의 특징 정보를 여러 개인과 연결시켜 개인 식별 방지

▶ 이름을 '김수철'(유명 가수 이름 등), 김삿갓(역사적 인물 등)으로 바꾸어 누군지 알 수 없도록 함

▶ 특정인의 몸무게를 20대 서울 거주 여성의 평균 몸무게로 처리하여 누구의 몸무게인지를 구분할 수 없도록 함

▶ 991202-1234567과 같은 주민번호를 99년생 남성으로 변환하여 개인을 식별할 수 없도록 함

▶ 이명박, 73세라는 특정인을 구분할 수 있는 경우에는 이씨 성을 가진 70대로 바꾸어 개인정보를 보호함

▶ 이명박, 안암1동 거주, 고려대학교 재학 이라는 특정인을 구분할 수 있는 경우에는 이○○, ○○대학 재학, ○○ 거주 식으로 처리함

2. 개인정보 비식별 조치 가이드라인

▣ 비식별화 기술 (1)

○ 일반적 기법 : 개인 식별요소 삭제 방법

처리기법	예시	세부기술
가명처리 (Pseudonymization)	ㅇ 홍길동, 35세, 서울 거주, 한국대 재학 → 임꺽정, 30대, 서울 거주, 국제대 재학	①휴리스틱 가명화 ②암호화 ③교환 방법
총계처리 (Aggregation)	ㅇ 임꺽정 180cm, 홍길동 170cm, 이콩쥐 160cm, 김팥쥐 150cm → 물리학과학생키합:660cm, 평균키 165cm	④총계처리 ⑤부분총계 ⑥라운딩 ⑦재배열
데이터 삭제 (Data Reduction)	ㅇ 주민등록번호 901206-1234567 → 90년대 생, 남자 ㅇ 개인과 관련된 날짜정보(합격일 등)는 연단위로 처리	⑧식별자 삭제 ⑨식별자 부분삭제 ⑩레코드 삭제 ⑪식별요소 전부삭제
데이터 범주화 (Data Suppression)	ㅇ 홍길동, 35세 → 홍씨, 30 ~ 40세	⑫감추기 ⑬랜덤 라운딩 ⑭범위 방법 ⑮제어 라운딩
데이터 마스킹 (Data Masking)	ㅇ 홍길동, 35세, 서울 거주, 한국대 재학 → 홍○○, 35세, 서울 거주, ○○대학 재학	⑯임의 잡음 추가 ⑰공백과 대체

2. 개인정보 비식별 조치 가이드라인

▣ 비식별화 기술 (2)

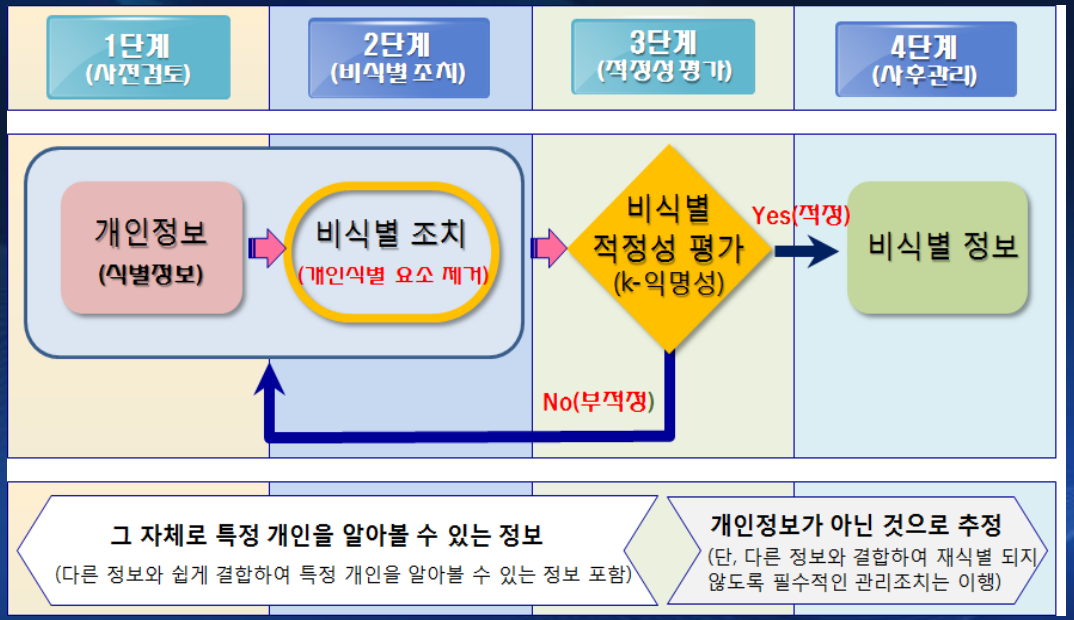
- **프라이버시 보호 모델 : 재식별 가능성 검토 기법**

기법	의미	적용례
k-익명성	특정인임을 추론할 수 있는지 여부를 검토, 일정 확률수준 이상 비식별 되도록 함	동일한 값을 가진 레코드를 k개 이상으로 함. 이 경우 특정 개인을 식별할 확률은 1/k임
l-다양성	특정인 추론이 안된다고 해도 민감한 정보의 다양성을 높여 추론 가능성을 낮추는 기법	각 레코드는 최소 l개 이상의 다양성을 가지도록 하여 동질성 또는 배경지식 등에 의한 추론 방지
t-근접성	l-다양성 뿐만 아니라, 민감한 정보의 분포를 낮추어 추론 가능성을 더욱 낮추는 기법	전체 데이터 집합의 정보 분포와 특정 정보의 분포 차이를 t이하로 하여 추론 방지

* k, l, t값은 전문가 등이 검토하여 마련

2. 개인정보 비식별 조치 가이드라인

▣ 단계별 조치사항



① 사전검토(개인정보 여부 검토)
↓
② 비식별조치(비식별 조치 기법 안내)
↓
③ 적정성 평가(외부전문가가 포함된 적정성 평가단*, K-익명성 모델** 활용)
↓
④ 사후관리(재식별방지 관리적·기술적 보호조치, 재식별정보 파기조치)

* 해당기관의 개인정보 보호책임자가 3명이상 관련 분야 전문가로 구성(단, 각 분야별 전문기관에서 운영하는 전문가 풀에서 외부전문가를 과반수 이상으로 위촉)
** 특정인에 대한 추론 여부를 검토, 그 가능성을 일정수준(1/k값) 이하로 낮추도록 함

2. 개인정보 비식별 조치 가이드라인

▣ 단계별 조치사항

⭘ 사전 검토 단계 : 개인정보 해당 여부 검토

○ 빅데이터 분석 등을 위해 정보를 처리하려는 사업자 등은 해당 정보가 개인정보인지 여부에 대해 아래 기준을 참조하여 판단

○ 해당 정보가 개인정보에 해당하지 않는 것이 명백한 경우에는 별도 조치 없이 빅데이터 분석 등에 활용 가능

⇨ 개인정보에 해당한다고 판단되는 경우 다음 단계의 조치 필요

2. 개인정보 비식별 조치 가이드라인

▣ 단계별 조치사항

○ 비식별 조치 단계 : 비식별 조치기법 적용

식별자(Identifier) 조치 기준

○ 정보집합물에 포함된 식별자*는 원칙적으로 삭제 조치

* '식별자'란 개인 또는 개인과 관련한 사물에 고유하게 부여된 값 또는 이름

○ 다만, 데이터 이용 목적상 반드시 필요한 식별자는 비식별 조치 후 활용

속성자(Attribute value) 조치 기준

○ 정보집합물에 포함된 속성자*도 데이터 이용 목적과 관련이 없는 경우에는 원칙적으로 삭제

* '속성자'란 개인과 관련된 정보로서 다른 정보와 쉽게 결합하는 경우 특정 개인을 알아볼 수도 있는 정보

- 데이터 이용 목적과 관련이 있는 속성자 중 식별요소가 있는 경우에는 가명처리, 총계처리 등의 기법을 활용하여 비식별 조치

○ 희귀병명, 희귀경력 등의 속성자는 구체적인 상황에 따라 개인 식별 가능성이 매우 높으므로 엄격한 비식별 조치 필요

2. 개인정보 비식별 조치 가이드라인

▣ 단계별 조치사항

⭘ 비식별 조치 단계 : 비식별 조치기법 적용

비식별 조치 방법

○ 가명처리, 총계처리, 데이터 삭제, 데이터 범주화, 데이터 마스킹 등 여러 가지 기법을 단독 또는 복합적으로 활용

※ '가명처리' 기법만 단독 활용된 경우는 충분한 비식별 조치로 보기 어려움

○ 각각의 기법에는 이를 구현할 수 있는 다양한 세부기술이 있으며, 데이터 이용 목적과 기법별 장·단점 등을 고려하여 적절한 기법·세부기술을 선택·활용

⇨ 비식별 조치가 완료되면 다음 단계의 조치 필요

2. 개인정보 비식별 조치 가이드라인

▣ 단계별 조치사항

○ 적정성 평가 단계 : k-익명성 모델 활용

적정성 평가 절차

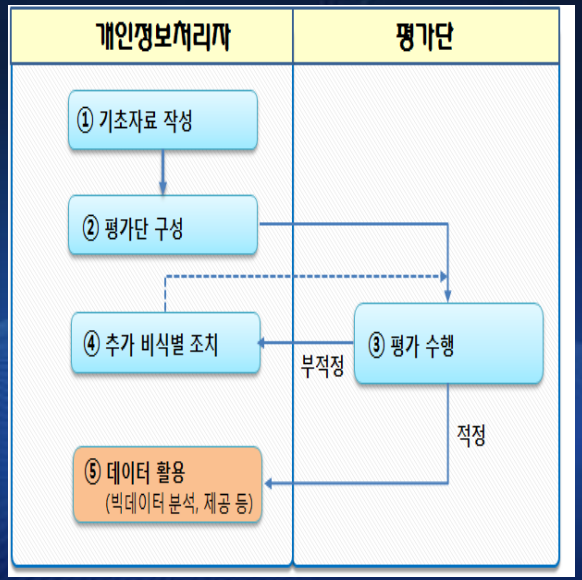
① (기초자료 작성) 개인정보처리자는 적정성 평가에 필요한 데이터 명세, 비식별 조치현황, 이용기관의 관리 수준 등 기초자료 작성

② (평가단 구성) 개인정보 보호책임자가 3명 이상으로 평가단을 구성(외부전문가는 과반수 이상: 최소 3명)

③ (평가 수행) 평가단은 개인정보처리자가 작성한 기초자료와 k-익명성 모델을 활용하여 비식별 조치 수준의 적정성을 평가

④ (추가 비식별 조치) 개인정보처리자는 평가결과가 '부적정'인 경우 평가단의 의견을 반영하여 추가적인 비식별 조치 수행

⑤ (데이터 활용) 비식별 조치가 적정하다고 평가받은 경우에는 빅데이터 분석 등에 이용 또는 제공이 허용



2. 개인정보 비식별 조치 가이드라인

▣ 단계별 조치사항

③ 적정성 평가 수행 : k-익명성 모델 활용

적정성 평가 수행

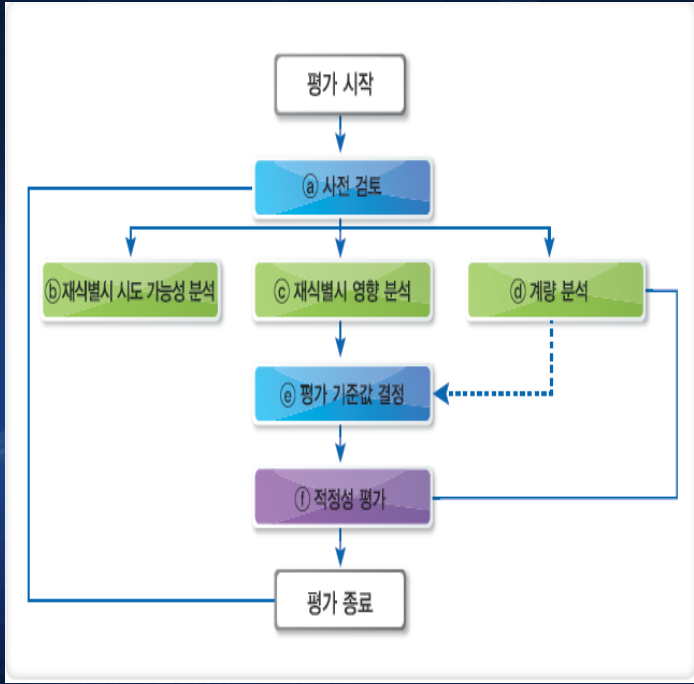
ⓐ (사전 검토) 개인정보처리자가 제출한 기초자료와 인터뷰등을 통해 평가대상 데이터의 개인 식별요소 포함 여부, 데이터 이용 목적, 비식별 조치 기범 등 검토

ⓑ (재식별 시도 가능성) 데이터를 이용 또는 제공받는 자의 재식별 의도와 능력, 개인정보 보호 수준등 재식별 시도 가능성 분석

ⓒ (재식별시 영향 분석) 데이터가 의도적 또는 비의도적으로 재식별될 경우 정보주체등에게 미칠 수 있는 영향 분석

ⓓ (계량 분석) 개인정보처리자가 제출한 K값의 정확성 여부 검증

ⓔ (평가기준값 결정) 평가단에서 '재식별 시도 가능성', '재식별시 영향','계량 분석' 결과와 데이터 이용 목적등을 종합적으로 고려하여 평가 기준값(K-익명성 값) 결정



2. 개인정보 비식별 조치 가이드라인

▣ 단계별 조치사항

○ 사후관리 단계

비식별 정보 안전 조치

○ 비식별 조치된 정보가 유출되는 경우 다른 정보와 결합하여 식별 될 우려가 있으므로 필수적인 보호조치 이행

- (관리적 보호조치) 비식별 정보파일에 대한 관리 담당자 지정, 대장관리, 이용 목적 달성시 파기 등의 조치가 필요함
- (기술적 보호조치) 비식별 정보파일에 대한 접근통제, 접속기록 관리 등의 조치 필요

○ 비식별 정보 유출 시 보호조치

- 유출 원인 파악 및 추가 유출 방지를 위한 관리적·기술적 보호조치
- 유출된 비식별 정보의 회수·파기

2. 개인정보 비식별 조치 가이드라인

▣ 단계별 조치사항

⭘ 사후관리 단계

재식별 가능성 모니터링

○ 비식별 정보를 이용하거나 제3자에게 제공하려는 사업자 등은 해당 정보의 재식별 가능성을 정기적으로 모니터링을 해야 함

○ 모니터링 결과, 점검 항목 중 어느 하나에 해당되는 경우에는 추가적인 비식별 조치 강구

○ 비식별 정보를 제공·위탁한 자가 재식별 가능성을 발견한 경우에는 이를 즉시 그 정보를 처리하고 있는 자에게 통지하고 처리 중단 요구 및 해당 정보를 회수 조치하여야 함

2. 개인정보 비식별 조치 가이드라인

▣ 단계별 조치사항

○ 사후관리 단계

비식별 정보 제공 및 위탁계약 시 준수사항

○ 비식별된 정보를 제3의 기관에 제공하거나, 처리 위탁하는 경우 재식별 위험관리에 관한 내용을 계약서에 포함

- (재식별 금지) 비식별 정보를 제공받거나 처리를 위탁 받은 사업자 등은 다른 정보와 결합을 통해 재식별 시도가 금지됨을 명시
- (재제공 또는 재위탁 제한) 비식별 정보를 제공하거나 처리를 위탁하는 자는 재제공 또는 재위탁 가능 범위를 정하여 계약서에 명시
- (재식별 위험 시 통지) 재식별이 되거나 재식별 가능성이 높아지는 상황이 발생한 경우에는 데이터 처리 중지 및 비식별 정보 제공자 또는 위탁자에게 통지 의무 등 명시

2. 개인정보 비식별 조치 가이드라인

▣ 비식별 조치 적정성 평가단 세부 평가수행 방법

1. 사전 검토
2. 재식별 시도 가능성 분석
   1) 재식별 의도 및 능력 분석
   2) 개인정보 보호 수준 분석
   3) 재식별 시도 가능성 분석
3. 재식별 시 영향 분석
4. 계량 분석
5. 평가 기준값 결정
6. 적정성 평가

2. 개인정보 비식별 조치 가이드라인

▣ 비식별화 조치 지원 및 관리 체계

○ **부처별로 지정한 전문기관을 통해 비식별화 지원 및 기관간 DB결합 지원**

- 미래부(정보화진흥원), 행자부·방통위(인터넷진흥원), 금융위(신용정보원), 복지부(사회보장정보원) 등
- 임시 대체키를 활용한 결합을 허용하는 경우에도 무분별한 결합을 통한 개인정보 침해 소지를 방지하기 위해 전문기관(제3의 공공기관)에서만 결합을 하도록 하는 등 지원 및 관리체계 필요

<분야별 전문기관의 역할>

- 비식별 조치 적정성 평가단 풀(비식별 조치 기법 전문가, 법률 전문가 등) 구성·운영
- 산업별로 필수적인 비식별 조치 이행 권고(k-익명성 수치 등)

  * 의료, 복지, 교육, 금융·신용, 통신, 유통, 공공·기타 분과
- 비식별 조치 적정성 실태 점검 등

2. 개인정보 비식별 조치 가이드라인

▣ 비식별화 조치 지원 및 관리 체계

○ **결합 절차**

① A社와 B社가 같은 알고리즘을 적용하여 식별자를 임시 대체키로 전환하고, 결합대상 정보집합물도 비식별 조치 및 적정성 평가 수행

※ '임시 대체키' 생성시 동 대체키에 잡음을 추가하거나, 2개 이상의 식별자를 활용할 경우 식별자 중 일부를 조합하여 불법적 복호화시에도 개인을 식별할 수 없도록 조치

② 비식별 조치된 정보를 전문기관에 제공 및 결합 요청

※ 이 경우 전문기관은 원본정보가 없어 제공받은 비식별 정보를 통해 특정 개인 식별 불가

③ 임시 대체키를 활용, 전문기관에서 결합 수행

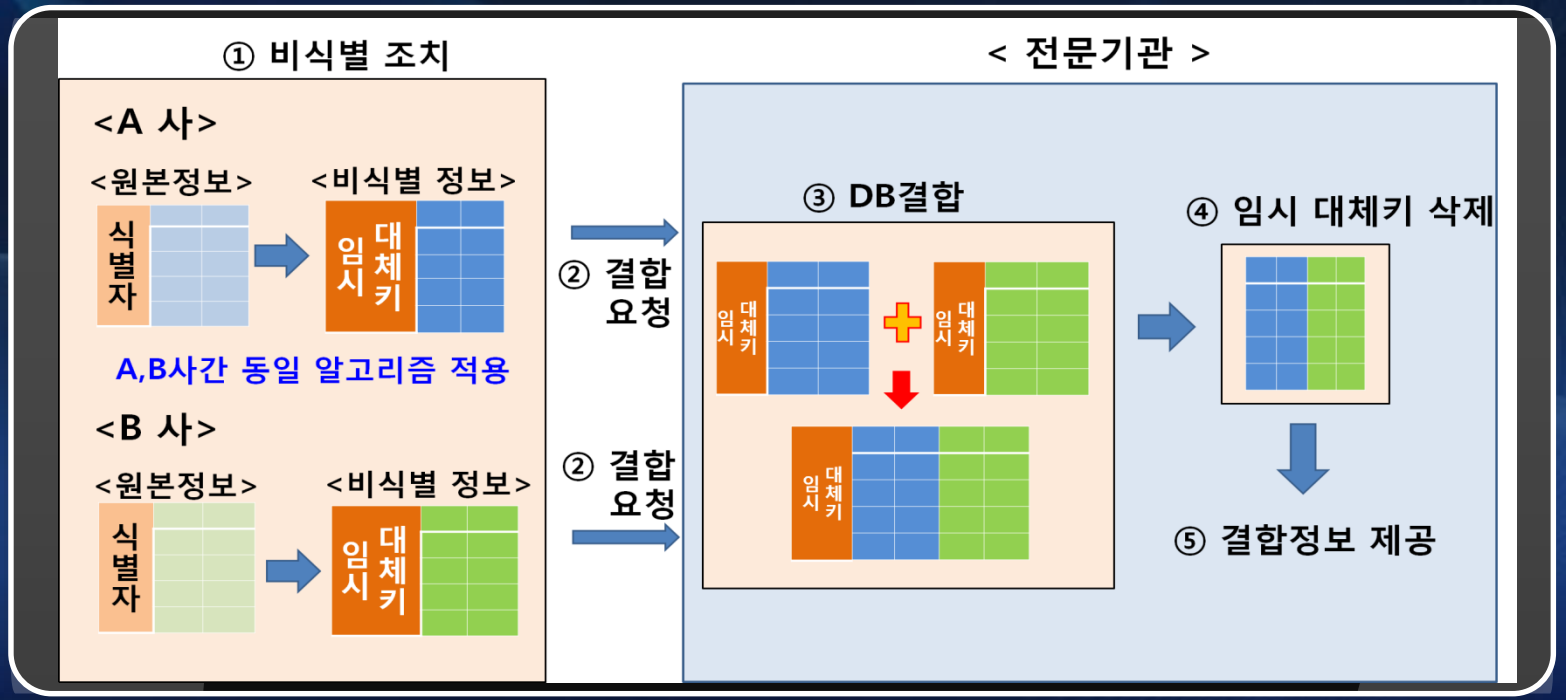
④ 임시 대체키 삭제

⑤ 결합 DB를 필요한 기업에게 제공(전문기관은 결합 후 파기 조치)

※ 임시 대체키가 삭제된 결합 DB가 제공되어 A와 B도 결합 DB를 통해 특정 개인의 식별이 어려움

2. 개인정보 비식별 조치 가이드라인

▣ 비식별화 조치 지원 및 관리 체계



2. 개인정보 비식별 조치 가이드라인

▣ 비식별화 조치 지원 및 관리 체계

○ 결합 시 유의사항

○ A와 B는 분야별 전문기관과 임시 대체키 생성 알고리즘에 대한 정보공유 금지

○ 임시 대체키 생성을 위해 주민등록번호를 활용하는 것은 금지

(개인정보 보호법 제24조의2, 주민등록번호 처리의 제한)

○ 다른 정보와의 결합을 위해 임시 대체키를 활용하는 경우, k-익명성 값은 임시 대체키를 제외하고 산출*

* 임시 대체키를 제외하지 않으면, 'k=1'로 산출되어 객관적 평가 불가

○ 전문기관은 결합 과정에서 재식별 발생시 해당 정보를 즉시 파기

○ 결합 DB를 제공받은 기관은 이용 전에 반드시 적정성 평가 수행

감 사 합 니 다

K-ICT 빅데이터센터