14. 해외 주요국 개인정보 가명·익명 처리 관련 법제 현황 분석

[목 차]

1. 개요

▶ 해외 주요국 중 EU의 GDPR과 독일, 영국, 미국, 일본의 개인정보 가명 처리 및 익명 처리 관련 법제와 활용 현황을 분석

▶ 가명 처리·익명 처리 기준 및 안전조치와 관련한 국제 표준으로서 ISO/IEC 20889 비식별 처리 기법에 대한 국제표준과 가명·익명 처리에 관한 주요국들의 지침 및 가이드라인을 분석

▶ 개인정보 2차 활용에 관련된 주요국의 법령 규정과 주요 활용 사례 검토

▶ 주요국 가명·익명 처리 활용 관련 감독기관 및 현황 검토

2. 해외 주요국 가명·익명 처리 관련 법령

가. EU

▶ 유럽연합 개인정보보호법(EU General Data Protection Regulation, 이하 GDPR))

- 가명 처리(pseudonymisation)는 추가 정보 없이는 더 이상 특정 개인정보 주체 정보 주체를 식별할 수 없는 방식으로 개인정보를 처리하는 것을 의미하며 추가적인 정보는 별도로 보관하고, 기술적·관리적 조치를 통해 해당 개인정보가 식별 가능한 개인에게 귀속되지 않도록 해야 한다고 정의하고 있음(제4조 제5항)
- 익명 정보(anonymisation information)란 정보주체를 더 이상 식별할 수 없도록 익명으로 처리된 정보"라고 정의하고 익명정보에는 GDPR이 더 이상 적용되지 않는다고 규정하고 있음(전문 제26항)
- GDPR 전문 제26항에서는 가명 처리된 정보도 식별가능한 개인에 관한 정보로 간주되어 개인정보로 취급됨을 명시하고 있음
- 그러나, GDPR만으로는 가명 처리와 익명 처리의 구체적인 구분 기준은 명시하고 있지 않음

▶ 안전조치로서의 가명 처리

- GDPR 제25조(설계 및 기본설정에 의한 개인정보보호)에서는 가명 처리를 개인정보 보호를 위한 적절한 기술적·관리적 조치의 하나로 언급하고 있음

▶ 가명 처리를 통한 정보 주체 정보 주체의 명시적 동의 없는 2차 활용 가능성

- GDPR 제6조 제4항에 따르면, 컨트롤러는 개인정보를 수집한 당초 목적과 양립 가능한(Compatible) 범위 내에서 정보 주체 정보 주체의 동의 없이 목적 외로 처리할 수 있는데, 양립 가능성을 판단하는 요소로 가명 처리 등의 안전조치가 존재하는지 여부가 규정되어 가명 처리가 이루어진 경우 양립 가능성이 인정될 가능성이 커짐
- GDPR 제89조 제1항에서는 공익적 기록보존, 과학적 또는 역사적 연구 목적, 통계 목적의 처리에 대해 가명 처리 등 안전조치를 적용할 것을 규정하고 있는데, 이 경우 해당 목적을 위한 가명정보의 처리에 있어서 정보 주체 정보 주체의 권리(제15조 열람권, 제16조 정정권, 제18조 처리제한권 및 제21조 반대권 등)가 일부 제한될 수 있음
- 위와 같이 가명 처리에 의한 정보처리는 개인정보를 보호하기 위한 안전조치로서의 의무적 성격과 동시에 개인정보처리자의 개인정보 활용 가능성을 높이는 역할도 하게 됨.

▶ EU 회원국 개별 국가에 대한 적용

- 위와 같은 EU GDPR 상의 가명·익명 처리에 관한 규정은 EU 회원국 전체에 적용되고 있으며, 각국은 해당 규정을 기준으로 가명·익명 처리에 관한 구체적 기준을 마련하고 있음

나. 독일

▶ 독일 연방 개인정보보호법(Bundesdatenschutzgesetz, 이하 'BDSG')[1)]

- 가명 처리는 추가 정보를 별도로 보관하거나 해당 데이터가 정보를 처리하는 당사자에게 배정되지 않도록 보장하는 기술적 및 조직적 조치를 통하여 추가 없이는 데이터를 구체적인 당사자에게 배정할 수 없는 방식의 개인정보 처리를 의미한다고 규정(법 제46조 제5호)
- 기본적으로 개인정보 보호 일반에 관한 부분은 GDPR의 적용을 받도록 하여 가명·익명 처리에 관한 부분은 GDPR에 따르도록 함(제2장)
- 가명 처리를 통한 정보 주체 정보 주체의 명시적 동의 없는 2차 활용 가능성

1) https://www.gesetze-im-internet.de/bdsg_2018/BJNR209710017.html

- BDSG 제27조에서는 "특별 범주의 개인정보 처리는 과학적, 역사적 연구 목적 또는 통계 목적을 위하여 처리가 필요하고 처리에 대한 책임자의 이해관계가 처리의 배제에 대한 당사자의 이해관계보다 현저히 더 중요한 경우 이 목적에 대한 동의 없이도 허용된다"고 규정하고 이 경우 가명 처리 등의 보호조치를 취하도록 함
- 이 경우 연구 목적 또는 통계 목적에 비추어 가능한 즉시 익명 처리하여야 하나, 이것이 당사자의 정당한 이해관계에 반하는 경우에는 예외로 함

다. 영국

▶ 영국 일반 개인정보보호법(UK General Data Protection Regulation, 이하 'UK GDPR')[2)]

- UK GDPR은 EU GDPR에서 규정한 비식별 조치 또는 익명성에 대한 개념과 기준, 가명화의 개념을 특별한 수정 없이 그대로 수용함
- 비식별화된 개인 데이터의 식별 금지
  - UK GDPR 제171조(1)에서는 개인정보 비식별 처리 책임자의 동의 없이 고의 또는 무분별하게 비식별 처리된 개인정보를 재식별하는 행위는 위법하다고 규정함
  - 이에 따르면 데이터를 제공받은 주체가 해당 데이터를 재식별하려는 행위를 하면 벌금 등 금전적 제재의 대상이 됨

라. 미국

▶ 개인정보 법령에 대한 미국의 법체계

- 미국은 EU GDPR이나 한국의 개인정보보호법과 같은 포괄적인 연방 차원의 개인정보보호법이 없음. 대신 의료, 교육 등 각 분야별 개별법으로 개인정보 비식별 내용을 다루고 있음
  - 교육부의 가족교육권 및 개인정보보호법(Family Educational Rights and Privacy Act, 이하 'FERPA') 관련 규정

2) https://www.legislation.gov.uk/ukpga/2018/12/contents

- 보건복지부의 건강보험 이동성 및 책임법(Health Information Portability and Accountability Act, 이하 'HIPAA') 프라이버시 규칙
- 식품의약국(FDA)의 식품안전 현대화법(Food Safety Modernization Act, FSMA) 관련 규정

- **(비식별정보 개념 사용)** 미국은 가명정보와 익명정보를 구분하지 않고, "비식별정보(de-identified information)"라는 개념을 사용함. 비식별정보는 개인을 식별할 수 없거나 개인을 식별할 수 있다는 합리적 근거가 없는 정보를 의미함
- **(자유로운 활용 허용)** 비식별정보에 대해서는 민간 자율규제를 원칙으로 하여 사실상 자유로운 활용이 가능함

▶ 캘리포니아 개인정보보호법(CCPA/CPRA)[3)]

- 캘리포니아주는 '18년 미국 최초의 민간 분야 개인정보 일반법인 캘리포니아 소비자 개인정보보호법(California Consumer Privacy Act. CCPA)을 제정하고, '20. 11. 이를 일부 개정하여 '23. 1. 1. 부터 California Privacy Rights Act(CPRA)를 시행하고 '23. 7. 1. 부터 집행하고 있음
- 개인정보에는 익명화된 소비자정보나 집계된 소비자정보는 포함되지 않음(§1798.140.(v)(3))
- **(가명화 정의)** 가명화 또는 가명 처리란 추가정보를 사용하지 않고는 특정 소비자에게 더 이상 귀속될 수 없는 방식으로 개인정보를 처리하는 것을 의미하며, 추가 정보는 별도로 보관되고 식별 가능한 개인정보가 소비자에게 귀속되지 않도록 하는 기술적·조직적 조치가 적용됨(§1798.140.(aa))
  - 재식별을 방지하는 기술적 안전장치 적용
  - 재식별을 특정적으로 금지하는 절차 마련
  - 비식별 조치된 정보의 의도치 않은 공개를 예방하는 절차 마련
  - 재식별 행위의 시도를 하지 않을 것
- 연구를 위한 가명 처리 이후 소비자의 삭제권이 제한될 수 있음
  - '연구'는 과학적 분석, 체계적 연구 및 관찰을 의미하며, 여기에는 공공 또는 과학적

3) https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?division=3.&part=4.&lawCode=CIV&title=1.81.5

지식을 개발하도록 설계되고, 공중 보건 분야에서 공익을 위해 수행되는 연구를 포함하되 이에 국한되지 않는 모든 기타 적용 가능한 윤리 및 개인정보 보호법을 준수하는 기초연구 또는 응용연구가 포함됨(§1798.140.(ab))

- 사업체가 정보를 삭제하면 해당 연구를 완료하는 것이 불가능하거나 완료 능력이 심각하게 손상될 가능성이 있는 경우(§1798.105(d)(6))

▶ HIPAA[4] 상 비식별 조치

- 이 법률은 보건의료정보의 보호와 활용에 관한 규제를 포함하고 있고, 그 하위법령인 'HIPAA 프라이버시규칙(HIPAA Privacy Rule[5])'이 구체적인 지침을 제공하고 있음
- HIPAA의 적용대상이 되는 보건의료정보는 '개별적으로 식별가능한 보건의료정보(Individually Identifiable Health Information)'임. 따라서 개별적으로 식별할 수 없는 보건의료정보는 HIPAA의 보호 대상이 아님
- HIPAA 프라이버시 규칙에서는 비식별 조치의 적용 기준과 구체적인 방법을 제시하고 있음
  - 합리적 기대가능성에 따라, 비식별 조치의 방법으로는 전문가 판단 방식과 18가지 식별자를 제거하는 방법(safe-harbor)을 정함

마. 일본

▶ 일본 개인정보 보호법[6][7]

- 익명가공정보
  - 개인정보의 구분에 따라 정하는 조치를 취하여 특정 개인을 식별할 수 없도록 개인정보를 가공해서 얻어지는 개인에 관한 정보로서, 해당 개인정보를 복원하는 것이 불가능하도록 한 것을 말함

· 해당 정보에 포함되는 성명, 생년월일 그 외의 기술 등을 통하여 특정 개인을 식별할 수

4) https://aspe.hhs.gov/reports/health-insurance-portability-accountability-act-1996
5) https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html
6) https://www.ppc.go.jp/personalinfo/minaoshi
7) 일본 개인정보 보호법에서는 '익명가공정보'와 '가명가공정보'라는 개념을 정의하여 비식별화 된 정보에 대한 규율을 하고 있음

있는 개인정보는 해당 개인정보에 포함되어 있는 성명, 생년월일 그 외의 기술 등 일부를 삭제하여 더 이상 개인정보를 복원하지 못하도록 하는 것(해당 일부의 기술 등을 복원할 수 있는 규칙성을 갖지 않는 방법을 통하여 다른 기술 등으로 대체하는 것을 포함)

· 개인 식별 부호[8]가 포함된 것은 해당 개인정보에 포함되는 개인 식별 부호의 전부를 삭제해야 함

• 가명가공정보

- 가명가공정보란 개인정보의 구분에 따라 정하는 조치를 취하여 다른 정보와 대조하지 않는 한 특정한 개인을 식별할 수 없도록 개인정보를 가공하여 얻은 개인에 관한 정보를 말함

· 해당정보에 포함되는 성명, 생년월일 그 외의 기술 등에 의해 특정의 개인을 식별할 수 있는 개인정보의 경우는 해당 개인정보에 포함되는 성명, 생년월일 그 외의 기술 등의 일부를 삭제하는 것

· 개인 식별 부호[9]가 포함된 것은 해당 개인정보에 포함되는 개인 식별 부호의 전부를 삭제해야 함

• 익명·가명가공정보의 처리에 관한 규정

	가명가공정보	익명가공정보
가공에 관한 규율	·규칙 제31조에 정하는 가공 기준에 따른 가공(법 제41조 제1항)	·규칙 제34조에 정하는 가공 기준에 따른 가공(법 제43조 제1항)

8) 개인 식별 부호란 다음 각 호 중 어느 하나에 해당하는 문자, 번호, 기호 및 기타 부호 중 법령으로 정한 것을 말함(법 제2조 제2항).
1. 특정개인의 신체 일부의 특징을 전자기기용으로 제공하기 위해 변환한 분자, 번호, 기호 및 기타 부호로서 그 특정 개인을 식별할 수 있는 것
2. 개개인에게 제공되는 서비스의 이용 또는 개인에게 판매되는 상품 구입에 대해 할당되거나 또는 개인에게 발급되는 카드 및 기타 서류에 기재되거나 전자적 방식으로 기록된 문자, 번호, 기호 및 기타 부호로서 그 이용자나 구입자 또는 발급받는 사람이 서로 다를 수 있도록 할당되거나 기재되거나 혹은 기록되어 특정 이용자나 구입자 또는 발급받은 사람을 식별할 수 있는 것

9) 개인 식별 부호란 다음 각 호 중 어느 하나에 해당하는 문자, 번호, 기호 및 기타 부호 중 법령으로 정한 것을 말함(법 제2조 제2항).
1. 특정개인의 신체 일부의 특징을 전자기기용으로 제공하기 위해 변환한 분자, 번호, 기호 및 기타 부호로서 그 특정 개인을 식별할 수 있는 것
2. 개개인에게 제공되는 서비스의 이용 또는 개인에게 판매되는 상품 구입에 대해 할당되거나 또는 개인에게 발급되는 카드 및 기타 서류에 기재되거나 전자적 방식으로 기록된 문자, 번호, 기호 및 기타 부호로서 그 이용자나 구입자 또는 발급받는 사람이 서로 다를 수 있도록 할당되거나 기재되거나 혹은 기록되어 특정 이용자나 구입자 또는 발급받은 사람을 식별할 수 있는 것

안전관리에 관한 규율	·삭제 정보 등의 안전관리 조치(법 제41조 제2항) ·가명가공정보의 안전관리조치(법 제23조, 제43조 제3항)	·가공 방법 등 정보의 안전 관리 조치(법 제43조 제2항) ·익명가공정보의 안전관리조치(노력의무)(법 제43조 제6항, 제46조)
작성시의 공표에 관한 규율	·이용 목적의 공표(법 제41조 제4항) (이용 목적을 변경하는 경우에는, 변경 후의 이용 목적에 대해서 공표 의무 있음)	·익명가공정보에 포함되는 개인에 관한 정보 항목의 공표(법 제43조 제3항)
제공에 관한 규율	·제3자 제공의 원칙 금지(법 제41조 제6항, 제42조 제1항, 제2항) ※법령에 근거하는 경우 또는 위탁, 사업 승계 혹은 공동 이용에 의한 예외 있음	·본인 동의 없이 제3자 제공 가능 ·제공시에 익명가공정보에 포함되는 개인에 관한 정보의 항목 및 그 제공 방법의 공표 및 익명가공정보인 취지를 제공처에게 명시 (법 제43조 제4항, 제44조)
이용에 관한 규율	·식별행위의 금지(법 제41조 제7항, 제42조 제3항) ·본인에의 연락 등의 금지(법 제41조 제8항, 제42조 제3항) ·이용 목적의 제한(법 제41조 제3항) ※이용 목적의 변경은 가능(법 제41조 제9항) ·이용 목적 달성시의 소거(노력의무)(법 제41조 제5항)	·식별행위의 금지(법 제43조 제5항, 제45조) ·불만처리(노력의무)(법 제43조 제6항, 제46조)

출처 : 일본 개인정보보호법 가이드라인, 가명가공정보와 익명가공정보의 취급에 관한 주된 규율의 차이

3. 가명·익명 처리 관련 국제 표준 및 주요국 지침·가이드

가. 국제 표준

▶ ISO/IEC 20889 비식별처리 기법에 대한 국제표준[10]

- ISO/ICE 20889 인증은 개인 데이터의 비식별화 기술에 관한 국제 표준을 정의함
  - 해당 인증은 데이터를 비식별화하는 과정에서 개인을 식별할 수 없도록 보장하는 방법과 절차를 기준으로 함
- **(주요내용)** 해당 인증은 비식별화 기술에 대한 용어 정리, 특성에 따른 비식별화 기술 분류, 재식별 위험을 줄이기 위한 기술의 적용 가능성 평가 등의 내용을 담고 있음
- **(비식별처리 기법 분류)** 각 기업별로 정의 및 작동원리, 장단점 분석, 적용 가능한 데이터 유형, 재식별 위험 평가 방법 등을 서술

- 통계적 기법 : k-익명성, l-다양성, t-근접성
- 암호화 기법 : 결정적 암호화, 순서 보존 암호화 등
- 삭제 기법 : 레코드 삭제, 속성 삭제 등
  - 일반화 기법 : 라운딩, 상·하단 코딩, 로컬 일반화 등
  - 가명화 기법 : 토큰화, 해싱 등
  - 무작위화 기법 : 노이즈 추가, 순서 변경 등

- **(비식별처리 프로세스)**
  - 식별자 분류
  - 비식별 기법 선택
  - 비식별 처리 수행
  - 유용성 및 프라이버시 평가
- 이외에 재식별 위험 평가 방법론, 비식별 데이터의 활용 시나리오, 비식별처리의 한계 및 주의사항에 관한 지침을 제공하고 있음

10) https://www.iso.org/standard/69373.html

나. 주요국 가명·익명 처리 지침

(1) EU

▶ 유럽네트워크 정보보안청(ENISA)의 「가명화 기술에 관한 지침」(Guidelines on Pseudonymisation Techniques)(2019. 11.)[11]

- '04년 설립된 유럽네트워크 정보보안청(European Union Agency for Cybersecurity, ENISA)은 유럽연합의 공식기관으로 유럽 전역의 사이버 보안수준을 높이는 것을 목표로 설립됨
- 해당 지침 주요 내용[12]
  - 가명화 기술을 구현할 때는 위험기반접근법을 채택
    · 필요한 보호 수준 평가
    · 유용성과 확정성 요구 사항 고려
    · 데이터 처리의 목적과 맥락 평가
  - 가명화 기술 : ▲카운터 ▲난수 생성기 ▲암호화 해시 함수 ▲메시지 인증 코드 ▲암호화
  - 가명화 정책 : ▲결정론적 가명화(Deterministic pseudonymisation) ▲문서-무작위화 가명화(Document-randomized pseudonymisation) ▲완전 무작위화 가명화(Fully-randomized pseudonymisation)
  - 처리의 위험성이 높은 경우 고급 가명화 시나리오 적용 : ▲비대칭 암호화 ▲더 정교한 키 관리 시스템 ▲여러 기술의 조합
- 의료

  ENISA가 '21. 1. 발표한 보고서인 <데이터 가명화: 고급기술 및 사용 사례(Data Pseudonymisation: Advanced Techniques and Use Cases)>[13]는 처리 과정에서 건강 데이터를 보호하기 위해 가명화 지침 제공

11) https://www.enisa.europa.eu/publications/pseudonymisation-techniques-and-best-practices/@@download/fullReport
12) 해외 개인정보보호 동향 보고서 2020. 1. 한국인터넷진흥원「ENISA의 가명 처리 기법 및 활용사례 보고서 분석」 참조
13) https://www.enisa.europa.eu/news/enisa-news/cybersecurity-to-the-rescue-pseudonymisation-for-personal-data-protection

(2) 영국

▶ 익명화, 가명화 및 개인정보 보호 강화 지침초안(Draft anonymistion, pseudonymisation and privacy enhancing technologies guidance)[14]

- 해당 지침 초안은 ▲익명화 ▲익명화의 효과성 ▲가명화 ▲책임성과 거버넌스 ▲개인정보 보호 강화기술(PETs) 등 총 5장으로 구성되어 있음
- ICO는 2022년 말 피드백을 반영하여 하나였던 지침을 ▲익명화 및 가명화 지침 ▲개인정보보호 강화기술(PETs) 지침의 두 개의 별도 문서로 나누어 발간하기로 결정
- PETs 지침은 '23. 6. 19. 최종 지침이 발표됨
- **(지침초안 주요 내용)** 해당 지침은 조직이 개인정보를 보호하면서 데이터의 가치를 활용할 수 있도록 가명화 기술을 올바른 이해와 구현을 돕는 것을 목적으로 함
  - 익명화된 데이터는 개인 식별이 불가능하지만, 가명화 된 데이터는 추가 정보를 통해 개인 식별이 가능
  - 가명화는 개인정보 처리 위험을 감소시키며, 데이터 보호 및 보안을 강화함
  - 가명화는 다른 목적으로 데이터 활용을 하는 것을 지원함
  - 가명화를 진행하기 위해서는 ▲목표 정의 ▲위험평가 ▲적절한 기술 선택 ▲결과 문서화의 과정을 거칠 것을 제안

▶ 영국 익명화 네트워크(UK Anonymisation Network, 이하 'UKAN') 보고서[15]

- UKAN은 2012년 설립된 비영리기관으로 여러 분야 전문가들을 통하여 익명화 사례를 연구하는 분석하는 활동을 수행하고 있음
- 2016년에 발표된 UKAN 보고서는 익명화 및 비식별화 기술의 중요성을 강조하며, 데이터 활용의 가능성과 그에 따른 재식별 위험에 대한 논의를 포함하고 있음

---

14) https://ico.org.uk/media/about-the-ico/consultations/4019579/chapter-3-anonymisation-guidance.pdf
15) 2018. 10.「해외 비식별 조치 가이드라인 등에 대한 비교·분석」, 한국인터넷진흥원. 104면 내지 109면

• 주요 내용

- 비식별 조치를 개인의 이름, 주소와 같은 직접 식별자를 제거하는 절차로 봄
  - 반면, 익명화는 특정 데이터에서 어떤 개인이 식별될 위험성을 무시할 수 있는 수준으로 낮추는 과정이라고 정의함
  - 이에 따라 비식별처리는 직접식별자를 제거하는 절차로 좁게 해석하고, 익명화는 직접 식별자의 제거를 포함해서 재식별의 위험성을 통제하는 접근법이란 좀 더 넓은 개념으로 봄
  - 위와 같은 개념 구분에 따라 가명화를 비식별 조치의 한 가지 기법으로 간주함
  - 데이터 상황적 접근법(data situation approach) : 데이터의 맥락에 따라 비식별화 방법이 달라져야 한다는 점을 강조하는 것으로 데이터가 사용되는 환경과 목적에 따라 적절한 비식별화 기술을 선택해야 한다는 것임
  - 재식별 위험성 관리 : 데이터를 공유하거나 활용하려는 주체에게 다음과 같은 세 가지 방법을 실천할 것을 강하게 권고함
    · 통계학자와 컴퓨터 분야 전문가들이 최고의 방법을 상용해서 적절한 수준의 기술적, 물리적, 관리적 보안 조치들을 평가할 것
    · 데이터의 유용성을 극대화하면서 재식별의 위험성을 최소화할 것
    · 개별 재식별 문제에 대응할 구체적 전략을 세울 것

(3) 아일랜드

▶ 아일랜드 개인정보 감독 기구 Data protection Commission(DPC)의 익명화 및 가명화에 관한 가이드라인(Guidance on Anonymisation and Psedudonymisation 2019)[16]

• 익명화와 가명화의 정의
  - 익명화 : 데이터를 비가역적으로 처리하여 개인을 식별할 수 없도록 하는 것. 익명화된 데이터는 더 이상 개인정보로 간주되지 않으며, 데이터 보호 법규의 적용을 받지 않음

16) https://www.dataprotection.ie/sites/default/files/uploads/2022-04/Anonymisation%20and%20Pseudonymisation%20-%20latest%20April%202022.pdf

- 가명화 : 개인 데이터의 식별자를 대체하여 직접적으로 식별할 수 없도록 하는 것. 가명화 된 데이터는 여전히 개인정보로 간주되며, GDPR의 적용을 받음

• 익명화 및 가명화의 용도

- 익명화된 데이터는 무기한 보관할 수 있으며 원래 목적 이외의 용도로 사용할 수 있음
- 가명화는 정보 주체 정보 주체의 신원을 제한적으로 보호하지만, 여전히 재식별 위험이 존재함

• 신원 확인 및 식별 가능성

- 데이터가 충분히 익명화되었는지 판단하기 위해서는 데이터 주체를 식별할 수 있는 모든 방법을 고려해야 함
- 직접적인 식별자가 제거되어도 다른 정보와 결합하여 개인이 식별될 수 있는 가능성이 있음

• 위험요소

- 재식별 위험은 여러 요인에 따라 달라지며, 특히 다른 데이터셋과의 연결 가능성이 중요함
- 개인적인 지식이 있는 경우 재식별 위험이 증가할 수 있음

• 익명화 기술

- 무작위화 : 데이터에 노이즈를 추가하여 개인과의 연결 고리를 끊음
- 일반화 : 데이터를 덜 세분된 형태로 변환하여 특정 개인을 식별하게 어렵게 만듦
- 마스킹 : 직접적인 식별자를 제거하여 개인 식별 위험을 줄임

• 데이터 보존 및 접근

- 익명화된 데이터는 개인 데이터를 포함하지 않으므로 보관기간에 대한 제한이 없음
- 원본 데이터가 남아 있는 경우, 해당 데이터는 여전히 개인정보로 간주됨

(4) 미국

▶ 미국 상무부(US Department of Commerce) 산하 국가기술표준원(National Institute of Standards and Technology, NIST) 보고서

- NIST는 미국 상무부 산하 기술의 표준화를 담당하는 기구인데, 비식별 조치에 대한 보고서를 발간함
- 비식별조치란 데이터셋에서 식별 정보를 제거하여 개별 데이터를 특정 개인과 연결할 수 없게 만드는 과정이라고 함
- 개인데이터의 비식별 조치(De-Indentification of Personal Data, 2015) 보고서[17)]
  - 비식별화는 개인정보의 수집, 처리, 보관, 배포 또는 출판과 관련된 프라이버시 위험을 줄일 수 있음
  - 최근 연구에 의하면 일부 비식별화된 데이터가 때때로 재식별 될 수 있음을 확인함
  - 재식별 위험에 대처하기 위해 데이터공격자에 대한 분석이 우선되어야 함을 강조
- 정부 데이터셋의 비식별화 : 기술 및 거버넌스(De-Identifying Government Datasets : Techniques and Governance, 2023)[18)]
  - '23. 9. SP 800-188 "정부 데이터셋 비식별화: 기술 및 거버넌스"는 NIST와 미국 인구조사국에서 작성한 것으로 정부 기관을 위한 비식별화 기술에 대한 지침을 제공하고 있음
  - 최종 버전은 '23. 9. 에 발표되었으며, 특히 차분 프라이버시 이론[19)]과 실제의 발전을 반영하여 이전 초안 버전을 업데이트함
  - 정부 데이터의 수집, 처리, 보관, 배포 또는 발행과 관련된 개인정보 위험을 줄이기 위해 비식별화를 사용하는 방법에 대한 구체적인 지침을 정부기관에 제공
  - ▲비식별화 기술 ▲비식별화 프로세스의 거버넌스 ▲연구자들에게 민감한 데이터셋을 제공하는 접근법 ▲공공투명성을 보장하는 방법 등을 내용으로 담고 있음

17) https://nvlpubs.nist.gov/nistpubs/ir/2015/nist.ir.8053.pdf
18) https://csrc.nist.gov/pubs/sp/800/188/final
19) '차분 프라이버시 이론'(Differential Privacy, DP)이란 개인정보를 보호하면서도 데이터의 통계적 특성을 유지하는 수학적 프레임워크임

(5) 일본

▶ 일본 개인정보의 보호에 관한 법에 대한 가이드라인(가명·익명가공정보 편)[20]

- 가명정보의 적정한 가공

- 개인정보에 포함되는 특정의 개인을 식별할 수 있는 기술 등의 전부 또는 일부를 삭제하는 것(해당 전부 또는 일부의 기술 등을 복원할 수 있는 규칙성을 갖지 않는 방법을 사용하여 다른 기술 등으로 대체하는 것을 포함함)

  - 개인정보에 포함되어있는 개인 식별 부호의 전부를 삭제하는 것(해당 개인 식별 부호를 복원할 수 있는 규칙성을 갖지 않는 방법을 사용하여 다른 기술 등으로 치환하는 것을 포함함)

  - 개인정보에 포함되어있는 것으로 이것이 부정하게 이용됨으로써 재산상 피해가 발생할 우려가 있는 기술 등을 삭제하는 것(예 : 신용카드 번호 삭제, 송금이나 결제 기능이 있는 웹 서비스의 로그인 ID·패스워드 삭제)

- 삭제 정보 등[21]의 안전관리 조치(규칙 제32조)

  - 법 제41조 제2항에 규정된 삭제정보 등(법 제 41조 제1항의 규정에 의하여 가명처리된 가공방법에 관한 정보에 있어서는 그 정보를 이용하여 가명처리정보의 작성에 사용된 개인정보를 복원할 수 있는 것에 한함)을 취급하는 자의 권한 및 책임을 명확히 규정할 것

  - 삭제정보 등의 취급에 관한 규정을 정비하고, 해당 규정에 따라 삭제정보 등을 적절히 취급하며 그 상황에 대한 평가 결과에 근거해 필요한 조치를 강구할 것

  - 삭제정보 등을 취급하는 정당한 권한을 갖지 않는 자에 의한 삭제정보 등의 취급을 방지하기 위하여 적절한 조치를 취할 것

---

20) https://www.ppc.go.jp/personalinfo/legal/guidelines_anonymous/
21) 가명가공정보의 작성에 이용된 개인정보에서 삭제된 기술 등 및 개인식별부호 및 법 제41조 제1항에 의해 이루어진 가공방법에 관한 정보를 말함

강구해야 할 조치	구체적인 예
▲삭제 정보 등을 취급하는 사람의 권한 및 책임의 명확화(규칙 제32조 제1호)	·삭제정보 등의 안전 관리 조치를 강구하기 위한 조직 체제의 정비
▲삭제 정보 등의 취급에 관한 규정의 정비 및 해당 규정에 따른 삭제정보 등의 취급 상황의 평가 및 그 결과에 기초하여 개선을 도모하기 위해 필요한 조치 실시(규칙 제32조 제2호)	·삭제정보 등의 취급에 관한 규정 등의 정비와 이에 따른 운용 ·직원 교육 ·삭제정보 등의 취급 상황을 확인하는 수단의 정비 ·삭제정보 등의 취급 상황의 파악, 안전관리조치의 평가, 재검토 및 개선
▲삭제정보 등을 취급하는 정당한 권한을 갖지 않은 자에 의한 삭제정보 등의 취급을 방지하기 위하여 필요한 조치 실시(규칙 제32조 제3호)	·삭제정보 등을 취급할 권한이 없는 사람에 의한 열람 등의 방지 ·기기, 전자 매체 등의 도난의 방지 ·전자 매체 등을 운반하는 경우의 누설 방지 ·삭제정보 등의 삭제 및 기기, 전자 매체 등의 폐기 ·삭제정보 등에 대한 접근 제어 ·삭제정보 등에 대한 접근자의 식별 및 인증 ·외부로부터의 부정 접근 등의 방지 ·정보 시스템 사용에 따른 삭제정보 등의 유출 방지

출처 : 일본 개인정보보호법 가이드라인 별표1. 삭제정보 등의 안전관리에서 요구되는 조치의 구체적 예

• 익명가공정보의 적절한 가공

기법명	해 설
항목, 레코드, 셀 삭제	가공 대상이 되는 개인정보 데이터베이스 등에 포함되는 개인정보의 기술 등을 삭제(항목 삭제, 레코드 삭제, 셀 삭제)
일반화	가공 대상이 되는 정보에 포함되는 기술 등에 대해서, 상위 개념 혹은 수치로 치환하는 것(예: 「오이」를 「야채」로 치환
상단(하단) 코딩	가공 대상이 되는 개인정보 데이터베이스 등에 포함되는 수치에 대해서, 특히 크거나 작은 수치를 정리하는 것(예 : 80세 이상의 수치 데이터를 「80세 이상」이라고 하는 데이터로 정리)
마이크로 어그리게이션	가공 대상이 되는 개인정보 데이터베이스 등을 구성하는 개인정보를 그룹화한 후, 그룹의 대표적인 기술 등으로 치환하는 것
데이터 교환	가공 대상이 되는 개인정보 데이터베이스 등을 구성하는 개인정보 상호에 포함되는 기술 등을 확률적으로 교환하는 것
노이즈 부가	일정한 분포에 따른 난수적인 수치를 부가함으로써 다른 임의의 수치로 치환하는 것
데이터 생성	인공적인 합성 데이터를 작성해, 이것을 가공 대상이 되는 개인정보 데이터베이스 등에 포함시키는 것

출처 : 일본 개인정보 보호법 가이드라인 별표2. 익명가공정보의 가공에 관련된 수법 예

4. 산업별 개인정보 2차 활용 관련 법제 동향 및 주요 사례

가. 주요국 관련 법령

▶ EU GDPR

- GDPR은 제6조 1(f) 및 제89조에 따라 정보 주체의 동의 없이 연구 목적으로 개인 데이터를 처리할 수 있도록 허용함
- 특히 제89조는 적절한 안전장치가 구현되는 경우 과학적 목적을 위한 추가 처리가 데이터 수집의 원래 목적과 양립할 수 있다고 명시하고 있음
- 또한 추가 처리는 과학적 목적, 역사적 목적 또는 통계적 목적을 위한 경우 공공의 이익에 부합하는 것으로 이해된다는 점도 명시되어 있음

▶ 영국

- UK GDPR은 과학적, 역사적 또는 통계적 연구 목적으로 개인 데이터를 처리하는 경우 일정 의무를 면제하는 연구 면제 조항을 규정하고 있음
  - 구체적인 조건에는 연구 목적을 훼손하지 않으면서 연구 목적을 위한 처리의 필요성이 포함될 수 있음
- 영국의 디지털 경제법(Digital Economy Act 2017)은 제5부 제5장에서 '연구 목적의 공유'를 규정하고 공공기관이 보유한 정보를 연구 목적으로 다른 사람에게 제공될 수 있도록 함[22)]
  - 그 정보가 특정인을 식별하는 경우, 공개되기 전에 특정인의 신원이 해당 정보 내에서 식별되지 않도록, 그리고 그 정보로부터 특정인의 신원을 추론하는 것이 합리적으로 가능하지 않도록 처리되어야 함
  - 공개를 위한 그 정보의 처리에 관여하는 모든 사람은 특정인을 식별할 수 있는 우발적인 정보의 공개 위험성을 최소화하고, 그러한 정보의 의도적인 공개를 방지하기 위한 합리적인 조처를 해야 함

22) 개인정보위원회「데이터 연계·결합 지원제도 도입방안 연구」2017. 12. 제vi면

▶ 미국

- 특정 상황에서 동의 없이 건강 데이터를 사용할 수 있도록 허용하는 HIPAA, 경우에 따라 동의 없이 연구 목적으로 교육 기록을 공유할 수 있도록 허용하는 FERPA, 캠퍼스 안전 및 공개에 관한 의무를 명시한 Clery Act 등 다양한 규정에서 연구용 개인 데이터 처리에 대한 예외가 명시되어 있음
- 또한, 개인 데이터를 포함하는 연구는 연방기관의 광범위한 지침에 따라 특정 조건 하에서 개인정보 보호법에 따라 면제될 수 있음

▶ 일본

- 개인정보 취급사업자가 학술연구기관 등인 경우로 그 개인정보를 학술연구 용도로 제공할 목적으로 취급할 필요가 있을 때, 학술연구기관 등에 개인 데이터를 제공하는 경우로 학술연구기관 등이 그 개인 데이터를 학술연구 목적으로 취급할 필요가 있을 때 애초 이용 목적의 제한을 받지 아니하고(법 제18조 제3항), 민감정보의 처리 시 동의를 필요하지 아니함(법 제20조 제2항 제5호, 제6호)
- 개인정보 취급사업자가 학술연구기관 등인 경우로서 개인 데이터 제공이 학술목적의 성과 공개 또는 교수를 위해 불가피한 때, 개인정보 취급사업자가 학술연구 기관 등인 경우로 개인 데이터를 학술연구 목적으로 제공할 필요가 있을 때(공동 학술연구), 제3자가 학술연구기관 등인 경우로 제3자 개인 데이터를 학술연구 목적으로 취급할 필요가 있을 때 정보 주체 정보 주체의 동의 없이 제3자 제공이 가능함(법 제27조)
- 다만, 위 경우 모두 개인의 권익을 부당하게 침해할 우려가 없어야 함

나. 산업별 활용 형태

▶ 의료 산업

- 개인 건강 데이터는 익명화 또는 가명 처리되어 임상실험 및 의료 연구에 사용됨. 예를 들어 제약회사는 환자 데이터를 분석하여 치료 효과의 경향을 파악함. IBM Watson Health와 같은 회사는 방대한 환자 정보를 사용하여 AI를 학습시켜 이를 통해 질병 진단 및 치료 추천을 지원함

▶ 금융 서비스

- 사기 탐지 : 은행과 금융 기관은 거래 데이터를 활용하여 사기 행위를 탐지하는 모델을 개발함. 소비 패턴을 분석하여 비정상적인 거래를 플래그[23]하는 방식임
- 신용 평가 : 대출 기관은 개인의 금융 데이터를 사용하여 신용도를 평가함. 이는 전통적인 신용 기록뿐만 아니라 유틸리티 및 임대료 지불 이력과 같은 대체 데이터 소스를 포함함

▶ 마케팅 및 광고

- 맞춤형 광고 : Google이나 Meta와 같은 회사는 사용자 데이터를 수집하여 상세한 사용자 프로필을 생성하고 이 정보를 바탕으로 사용자 관심사와 행동에 기반한 맞춤형 광고를 제공하는 데 사용함
- 고객 인사이트(Customer Insight) : 소매업체는 구매 패턴과 인구통계정보를 분석하여 마케팅 전략을 맞춤화 함. 예를 들어 Amazon은 브라우징 기록과 구매 데이터를 사용하여 제품을 추천함

▶ 통신산업

- 네트워크 최적화 : 통신회사는 통화데이터 기록을 분석하여 네트워크 성능 및 고객 서비스를 개선함. 이 분석은 사용 패턴을 이해하고 자원 배분을 최적화하는 데 도움이 됨
- 이탈 예측 : 고객의 사용 패턴과 피드백을 분석하여 통신 제공업체는 어떤 고객이 서비스를 이탈할 가능성이 높은지를 예측하고 이를 바탕으로 유지 전략을 시행함

▶ 운송 및 물류

- 경로 최적화 : Uber와 같은 회사는 승차 데이터를 활용하여 운전자의 경로를 최적화하고, 고객의 대기 시간을 줄임
- 예측 유지보수 : 항공사는 비행 데이터를 분석하여 항공기의 유지보수 필요성을 예측함으로써 안전성을 높이고 가동 중단 시간을 줄임

▶ 전자상거래

- 개인화 : 전자상거래 플랫폼은 개인의 쇼핑 데이터를 활용하여 사용자 경험을 맞춤화

23) 데이터를 분석하여 비정상적인 패턴을 식별하는 것을 의미함

함. 예를 들어 Netflix는 시청 기록을 기반으로 개별 선호에 맞춘 프로그램과 영화를 추천함

- 재고관리 : 소매업체는 판매 데이터를 분석하여 제품 수요를 예측하고, 이를 통해 재고를 보다 효과적으로 관리

다. 주요 사례

▶ 독일 보험회사 Provinzial Rheinland(2021년) 사례[24)]

- 독일 보험회사 Provinzial Rheinland(이하 'Provinzial')는 이미 소비자의 종류별 보험 수요를 예측하고 그에 맞는 제품을 추천하는 'Next best offer' 모델을 보유하고 있었으나, 이 마케팅 모델을 강화하기 위해 재현데이터[25)]를 활용
- Provinzial은 데이터 가용성, 모델 사용, 개인정보보호 규정 세 가지 지표를 만족하는 재현데이터를 생성했고 성능을 테스트하는 과정을 진행하였고, 원데이터와 재현데이터를 비교한 결과 재현데이터의 80% 정도를 사용할 수 있음을 확인했으며, 재현데이터로 학습시킨 모델의 성능이 원본 성능의 97% 이상을 충족했다고 함
- 이러한 결과는 원데이터와 재현데이터를 적절히 섞어 활용해 개인정보 노출 우려 없이 기존 머신러닝 모델을 발전시킨 것에 의의가 있음

▶ 영국 SynAE project(2018~2023년)[26)]

- 영국의 국민보건서비스(National Health Service, 'NHS') Digital[27)]이 제공하는 SUS(Secondary Uses Service) 데이터에서 추출한 내원 및 응급 데이터와 입원 환자 치료 데이터를 결합·가공하여 재현데이터로 구현한 시범 사업임
- 환자의 개인정보 노출을 막으면서도 데이터 공유를 확대하기 위한 취지로 시작된 이 프로젝트는 영국 정부의 NHS England 의무조항을 따르도록 수행되었으며 원데이터에

24) 2023 연구보고서「데이터 가명·익명 처리 기법의 현황과 대안 : 재현데이터를 중심으로」김현태·장가영. 보험연구원, 제50면
25) 재현데이터란 관측된 원데이터를 생성하는 모집단을 가정하고, 통계적·기계학습모형을 통해 생성한 모의 데이터를 의미함
26) 2023 연구보고서「데이터 가명·익명 처리 기법의 현황과 대안 : 재현데이터를 중심으로」김현태·장가영. 보험연구원, 제48면
27) 영국은 국가가 의료보장을 하고 있어 보건의료정보가 NHS Digital이라는 기관에 집적됨

민감정보가 다수 포함되어 있어 국가통계청(Office for National Statistics, 'ONS')에서 이 프로젝트를 검수함

- 지역 인구통계학적 정보에 기반하여 지리적 정보를 제거하고 세분화된 연속형 변수를 밴드로 그룹화[28]하고, 입원 일시, 시각과 같이 정확한 시간정보를 제거하고 이상치 정보는 추출해 마스킹 처리. 추가로 고유한 값을 가진 관측치 및 일부 희귀 부분 집합도 제거하여 노출 위험을 줄임
  - 이 프로젝트에서 산출된 재현데이터는 매년 공개되고 건강·보건 관련 데이터 공개는 다양한 건강정보를 활용한 서비스 개발을 가능하게 하고 있음

▶ 영국 가족자원조사(Famaily Resource Survey, 'FRS')[29]

- FRS는 현재 영국 국민들의 생활 수준과 상황을 파악하기 위하여, 영국 내 개인 가구 대표 표본을 통해 개인들의 소득과 그 상황에 관한 정보를 매년 수집하는 조사임
- FRS의 주요 목표는 영국 노동연금부(Department for Work and Pension, 'DWP')에 사회복지 정책의 개발과 모니터링 및 평가를 위한 정보를 제공하는 것임
- FRS 데이터는 통계규제국(Office for Statistics Regulation, OSR)에 의해 국가 통제 대상으로 지정되어 있음
- FRS 데이터와 행정데이터의 연결을 통하여 다양한 금융에의 활용, 데이터 품질 제고, 응답자 및 사용자 이점을 실현할 수 있는 가능성을 높임

▶ 미국 SIPP Synthetic Beta('SSB')(2013~2023)[30]

- SSB는 SMS 가구 조사에서 비롯된 개별 관측치 수준의 통계기초자료(Microdata)[31]를 세금 및 사회보장 데이터와 통합한 결과물임

---

28) 연속적인 값을 가지는 변수를 일정한 구간으로 나누어 범주화 하는 것을 의미함
29) 한국보건사회연구원「보건복지 분야 데이터 경제 활성화를 위한 다출처 데이터 연계, 통합, 활용방안 연구(II)-가명정보 결합을 중심으로」2023. 제96면
30) 2023 연구보고서「데이터 가명·익명 처리 기법의 현황과 대안 : 재현데이터를 중심으로」김현태·장가영. 보험연구원, 제47면
31) 마이크로데이터는 통계조사의 원자료(Raw data)에서 개인정보, 입력오류, 논리오류 등을 수정한 개별 단위(개인, 가구, 사업체 등)의 자료를 의미함

- 원데이터는 Survey of Income and Program Participation 조사에 응한 응답자들의 설문기록이 담겨 있으며, 사회보장행정(SSA)/내부수입서비스(IRS) 양식 기록과 퇴직 및 장애 급여 수령에 대한 행정 기록 등 민감정보가 포함되어 있었음
- 노출위험을 줄이기 위해 미국 인구조사국(Census Bureau)은 2013년부터 매년 누적된 데이터를 통합하여 새로운 버전의 부분 재현데이터를 가공하여 배포함
- 소속 경제학자, 통계학자와 대학 연구원들이 협력하여 세금, 수입 등 개인정보가 담긴 데이터를 부분 재현하고, 데이터 구조의 보존을 위해 변수 간 관계를 유지하도록 하고 각 관측치 기록을 대체하는 방식으로 연구를 진행
- 데이터 공개 전 노출위험에 대한 심사를 진행하고 SSB의 기록과 외부 데이터를 연결해도 개인정보 파악이 불가능할 것이라는 판정을 받음
- SSB는 최종적으로 데이터 공개 검토 위원회와 IRS, SSA 위원회의 승인을 받은 후 서버에 안전하게 등록됨
- 현재 SSB는 코넬 대학교의 가상 연구 데이터 센터에 있는 SDS(Synthetic Data Server)에 저장되어 있으며, 연구자는 서버에서 무료 계정을 만들고 키를 얻으면 SSB데이터를 사용할 수 있음
- SSB는 주요 변수에 대한 통계적 특성이 원데이터와 매우 유사하다고 알려져 있고 가장 광범위하게 공개된 재현데이터로 평가받음

▶ 일본 건강보험 청구데이터와 건강검진 데이터의 결합 데이터베이스의 활용[32)]

- 진료 보수 명세서 정보·특정 건강진단 등 정보 데이터베이스(이하 'NDB')는 의료비 적정화 계획의 작성, 실시 및 평가를 위한 조사나 분석 등에 이용하는 데이터베이스임
- 일본은 「진료 보수 명세서 정보·특정 건강진단 등 정보에 관한 가이드라인」을 정비하여 NDB의 제3자 제공을 2011년부터 시범실시, 2013년부터는 본격적으로 실시함
- NDB에는 건강보험 청구 데이터와 건강검진 데이터가 익명화 처리를 거쳐 결합된 후 저장되어 있음
- 일본 의료제도의 특성으로 인해 전국 단위 데이터가 부재한 실정에서 지역 데이터를

32) 한국보건사회연구원 「보건복지 분야 데이터 경제 활성화를 위한 다출처 데이터 연계, 통합, 활용방안 연구(II)-가명정보 결합을 중심으로」 2023. 제115면

결합하여 전국 단위 데이터인 NDB로 구축하여 중앙부처 및 지자체에서는 의료비 적정화 계획의 작성 및 평가 등 정책적으로 활용할 수 있게 됨

- NDB는 보험자, 연구기관, 대학 및 국가 지원 연구를 수행하는 연구자(민간기업 포함) 등 제3자에게 제공될 수 있으며, 이들에 의해 의료서비스의 질 향상 등에 기여할 수 있는 분석 및 연구에 활발하게 활용되고 있음

5. 가명·익명 처리 활용 관련 주요국 감독기관 현황

▶ 영국

[NHS Digital]

- 영국은 국가보건서비스(National Health Service, 이하 'NHS')라는 공공 의료 서비스 시스템을 운영하고 있어서 OECD 국가 중 가장 광범한 국가적 보건의료 데이터셋을 보유하고 있음
- 잉글랜드 지역 보건의료 분야 데이터 연계는 '임상시험연구데이터링크(CPRD)'가 담당하고 있으며, 공공보건 연구를 위해 익명화된 1차 진료기록을 제공해 옴
  - 매년 CPRD는 공공보건연구 목적으로 익명화된 연계 데이터를 제공하는 것에 대해 보건연구당국의 Section 251 규제 승인을 받아야 함
  - 데이터 연계는 환자식별정보를 합법적으로 수집할 권한을 가진 잉글랜드 법정기구인 NHS Digital에 의해 이루어짐
- care.data 프로그램 논란
  - care.data는 일반의가 보유한 환자정보를 HSCIC의 국가 데이터베이스에 집적하여 NHS England의 사업으로 2013년에 시작되었음
  - care.data의 데이터는 직접적인 진료 목적이 아닌, 의료 서비스 기획이나 의학 연구 등 2차적 목적으로 위해 활용되고, NHS 외부의 승인된 제약회사, 보건 자선단체, 대학, 병원 위탁단체 및 다른 사기업 등에 제공될 수 있었음
  - 그러나, 정보주체의 동의 없이 민감한 의료정보가 HSCIC에 집적되고 연구자 등 제3자에게 제공됨으로써 정보주체 통제권 약화, 해킹 등을 통한 개인정보 유출, 보험회사 등 영리적인 목적을 위한 제공, 비식별 데이터의 재식별 가능성 등 개인정보 침해에 대한 우려가 제기되어, 큰 반발을 야기하게 되고, 많은 논란 끝에 2016년 7월 care.data 프로그램은 취소됨

[영국 ADRN]

- 영국의 행정데이터연구네트워크(ADRN)는 사회, 경제 연구자들에게 안전한 환경에서 연계된 비식별 행정 데이터를 제공하기 위한 네트워크임
  - ADRN은 연구자들을 대신하여 요청할 데이터의 범위를 검토하고, 데이터 보유기관과 협의를 진행하며, 신뢰할 수 있는 제3자를 통해 데이터를 연계하고, 연계된 비식별 데이터에 접근할 수 있는 보안 환경을 제공하는 기구임(직접 데이터를 보유하지는 않음)

▶ 미국

[미국 국가보건통계센터]

- 보건의료서비스를 민간 제공자와 보험사가 주도하고 있는 미국은 보건의료 데이터의 수집 및 보관도 다양하게 분산되어 있음
  - 따라서 연구 목적의 데이터 접근을 위한 계약 체결도 개별 업체를 매개로 해야 함
  - 민간 영역의 데이터 연계는 통상 업체 내에서 이루어지며, 다른 업체와의 데이터 연계에 제한을 두고 있음
- 미국 국가보건통계센터(NCHS)는 자신이 보유한 데이터의 식별자를 제거한 후 공개사용 데이터 파일로 만들어 공개하고 있음[33)]
  - 데이터셋은 통계 보고 및 분석 목적으로만 사용해야 함
  - 의도하지 않게 개인 및 기관의 신원이 노출된 경우 이를 사용하지 말고 NCHS 책임자에게 알려야 함
  - 이 데이터셋을 개인 식별이 가능한 다른 NCHS의 데이터나 외부 데이터와 연계해서는 안 됨

[미국 Data Linkage Infrastructure]

- 미국 인구조사국은 평가자 및 정책 분석가의 행정 데이터 접근을 증진하기 위해 데이터 연계기반(Data Linkage Infrastructure)을 확대해 옴

33) 미국 공공보건서비스법 Section 308(d)

- 연계기반의 데이터를 사용하기 위해서, 연구자는 제안서를 제출해야 하고, 이 프로젝트가 외부의 데이터를 연계기반으로 가져올 경우 그 외부 데이터의 이용과 전송을 허가하는 서신을 제출해야 함
- 연구 제안서에 대해서는 학술적인 가치, 실행 가능성, 잠재적인 공개 위험성 등이 평가됨
- 제안서의 승인이 완료되면, 연방조사국은 연방통계연구데이터센터로 접근을 허용함
- 연구자들은 승인된 데이터 파일의 읽기전용 비식별화 버전에 접근하게 됨

▶ 독일 GRLC/FDZ

[GRLC/FDZ]

- 독일레코드연계센터(GRLC)는 사회과학 분야의 학술 연구를 위해, 행정 데이터를 이용한 데이터 연계 활성화를 목적으로 설립됨
  - GRLC는 프라이버시를 보호하면서도 효과적으로 데이터 연계를 수행할 수 있는 방법에 대한 연구라 진행하고 있음
- 독일연방고용국 연구데이터센터(FDZ)는 정부프로젝트 연구자에게 사회보장 및 고용 분야에서 비영리적 실증 연구를 위해 마이크로데이터에 대한 접근을 제공함
  - 연구자의 데이터 접근을 위해 데이터의 익명화 정도 및 이용 조건에 따라 현장 이용, 원격 데이터 접근, 학술적 이용 파일 등 세 가지 방법을 제공함

▶ 개인정보 감독 기구가 승인하는 국가

- 아이슬란드와 덴마크 일부 국가에서는 개인정보 감독 기구가 연구신청서 승인 기구의 역할을 하고 있으며, 프랑스의 경우 보건 분야에서 공익적인 연구, 조사 혹은 평가를 목적으로 한 개인정보의 처리, 그 목적이 서로 다른 공익을 위한 파일들의 연계 등에 대하여 개인정보 감독 기구인 CNIL의 허가를 받고 있음

6. 시사점

▶ 익명·가명 처리에 대한 주요국의 법제 현황 분석을 통하여 가명 처리와 익명 처리는 개인정보를 보호하는 안전조치이자 개인정보를 유용하게 활용할 수 있게 하는 이중적인 기능이 있음을 확인할 수 있음

▶ GDPR 제정 이후 가명정보는 개인정보로서 데이터 보호법령의 규율의 대상이 된다는 점에서는 세계 주요국의 법령에서는 차이가 없으며, 이로 인하여 데이터 보호 법령이 적용되지 않는 익명정보와 가명정보의 구분 필요성이 확대되고 있음

- 이에 따라 각 국에서는 가명 처리와 익명 처리를 구분하는 지침과 가이드라인을 마련하였거나 만들어 나가고 있는 상황임

▶ 주요 국가들은 데이터 연계와 결합을 통하여 개인정보의 2차 활용을 통하여 의료, 복지 등 공익의 향상에 상당한 결과를 내는 사업을 지속적으로 펼치고 있음

- 특히 공공데이터에 있어 그 활용도가 높음을 알 수 있음
- 특정 기관이 데이터 연계와 결합을 위한 활용을 책임지고 있으며, 일부 국가에서는 개인정보 감독기관이 직접적인 감독의 역할을 하고 있음

▶ 익명 처리 및 가명 처리를 통한 데이터 활용은 특히 AI 산업 발전에 있어서 필수적인 조건에 해당하는바, 그 활용에 관한 기준 및 이에 대한 데이터 거버넌스 형성을 위한 규율은 계속해서 만들어질 것으로 예상됨

▶ 주요 국가들은 공익에 기여하면서도 안전한 데이터 거버넌스를 위해서 기술적인 조치뿐만 아니라 데이터의 이용과 보호에 관련된 법제, 데이터 접근·연계 정책, 연구기관 혹은 데이터 연계기관의 인증, 심사 절차, 데이터 접근 절차 등에 이르는 전반적인 보호 체계를 구축해 나가고 있음

- 이 모든 과정이 체계적으로 조율되지 못한다면, 자칫 데이터 연계 및 제공 과정에서 개인정보 침해가 발생하거나, 반대로 공익에 기여할 수 있는 데이터의 활용을 제약할 수 있을 것임
- 따라서 데이터 활용을 활성화하기 위해서는 데이터의 활용 및 보호를 위한 데이터 거버넌스체계 수립이 중요하다고 할 것임

▶ 기술의 발전으로 인하여 가명정보 및 익명정보에 대한 재식별 위험성은 더 커질 수밖에 없다는 세계 각국 보고서들의 발표는 비식별정보 활용에 있어 지속적인 연구가 필요함을 시사함

- 언급한 영국의 care.data 사건에서 보듯이 익명 처리·가명 처리를 통한 데이터 활용에 대한 장벽은 정보 주체 정보 주체가 신뢰할 수 있고, 투명성이 보장된 형태의 활용이 이루어져야 제거될 수 있으며, 그것이 전제되어야 데이터 활용이 지속적으로 확대될 수 있음을 중요하게 인식해야 할 것임

[참고 자료]

- EU General Data Protection Rugulation
- 독일 연방 개인정보보호법
- 영국 일반 개인정보보호법
- 캘리포니아 소비자 프라이버시법
- 보건의료정보의 이동과 책임에 관한 법률(HIPAA)
- 일본 개인정보 보호법
- ISO/IEC 20889 비식별처리 기법에 대한 국제표준
- 유럽네트워크 정보보안청(ENISA)의 「가명화 기술에 관한 지침」(Guidelines on Pseudonymistion Techniques)
- 해외 개인정보보호 동향 보고서, 한국인터넷진흥원 「ENISA의 가명 처리 기법 및 활용사례 보고서, 2020. 1.
- 영국 익명화 네트워크(UK Anonymisation Network) 보고서
- 「해외 비식별 조치 가이드라인 등에 대한 비교·분석」, 한국인터넷진흥원, 2018. 10.
- 아일랜드 개인정보 감독기구 Data protection Commission(DPC)의 익명화 및 가명화에 관한 가이드라인(Guidance on Anonymisation and Psedudonymisation 2019)
- 미국 국가기술표준원 개인데이터의 비식별 조치(De-Identifcation of Personal Data, 2015) 보고서
- 미국 국가기술표준원 정부 데이터셋의 비식별화 : 기술 및 거버넌스(De-Identifying Government Datasets : Techniques and Governance, 2023)
- 일본 개인정보의 보호에 관한 법에 대한 가이드라인(가명가공정보·익명가공정보 편), 2016. 11.(2024. 12. 일부개정)
- 개인정보위원회 「데이터 연계·결합 지원제도 도입방안 연구」 2017. 12.
- 2023 연구보고서 「데이터 가명·익명 처리 기법의 현황과 대안 : 재현데이터를 중심으로.」 김현태·장가영. 보험연구원
- 한국보건사회연구원 「보건복지 분야 데이터 경제 활성화를 위한 다출처 데이터 연계, 통합, 활용방안 연구(II)-가명정보 결합을 중심으로」 2023.