경북대학교 법학연구원『법학논고』
제79집 (2022.10) 67~99면.
Kyungpook Natl. Univ. Law Journal
Vol.79 (Oct 2022) pp.67~99.
http://doi.org/10.17248/knulaw..79.202210.67

가명정보에 관한 법적 통제

김 창 조*

〈국문초록〉

가명정보는 개인을 알아볼 수 있는 정보나 다른 정보와 쉽게 결합하여 알아볼 수 있는 개인정보를 동 정보의 일부를 삭제하거나 일부 또는 전부를 대체하는 등의 방법으로 추가 정보가 없이는 특정 개인을 알아볼 수 없도록 처리하는, 즉 가명처리를 함으로써 원래의 상태로 복원하기 위한 추가정보의 사용·결합 없이는 특정개인을 알아볼 수 없게 된 정보를 말한다.

주관적 공권으로서 우리 헌법 해석상 인정되어 온 개인정보자기결정권은 정보주체가 타인이 보유하고 있는 자신의 정보에 접근하여 열람하거나 그 정보의 정정, 삭제, 차단 등을 요구함으로써 자신에 관한 정보에 통제할 수 있는 권리이다. 이를 보장하기 위해서 개인정보보호법에서는 정보주체의 동의를 개인정보처리의 원칙적인 요건으로 정하고 있는데 일반적인 공익보다 개인정보자기결정권을 우선시키는 것으로 평가된다.

동의요건 준수를 위하여 많은 사회적 비용이 발생한다. 이러한 점을 고려하여 개인정보보호법은 개인정보자기결정권을 보호하면서 동의요건 완화를 통한 개인정보의 이용확대를 위해서 가명정도제도를 적극적으로 법상 규정하게 되었다. 빅데이터와 인공지능이 산업과 사회발전의 핵심을 구성하는 새로운 미래를 준비하기 위해서는 대량의 데이터 수집과 분석이 필수 불가결한 것이고 가명정보제도가 가장 현실적 대안이라고 할 수 있다. 앞으로 개인정보보호를 강화하면서 가명정보를 통한 데이터 활용범위가 확대되도록 제도운용의 균형을 유지할 필요가 있다. 이러한 것을 실현하기 위하여 다음의 3가지 점은 추후 개선이 필요하다.

첫째, 추가정보는 개인정보의 전부 또는 일부를 대체하는 처리과정에서 생성 또는 사용된 정보로서 특정 개인을 알아보기 위하여 사용·결합될 수 있는 정보이다. 이러한 추가정보에 대한 법상 정의규정을 두고 있지 않다. 추가정보 범위를 적정하게 설정하지 않을 경우, 가명정보 의 적정한 범위 획정 상 혼란을 가져올 수 있다. 추가정보의 개념을 명확히 함으로써 제도 시행의 역기능을 사전에 예방할 필요가 있다.

둘째, 가명정보에 규제에 관한 법령상의 근거를 두지만, 가명처리의 핵심적 절차는 가이드라인의 형태로 규율된다. 행정권의 발동에는 작용법적 수권을 요한다는 관점에서 보면 가명처리에 관한 규율은 법률에 의한 행정이 아닌 행정지도에 의한 행정으로 변질될 우려가 있다. 가명처리에 관한 법적 규율밀도를 높이는 것이 필요하다.

* 경북대학교 법학전문대학원 교수/법학연구원 연구위원

셋째, 데이터 결합제도에 대하여 과도하게 우리 개인정보보호법은 규정기반 접근법(rule based approach)을 취하고 있다. 부분적으로 위험기반 접근법(risk based approach)을 도입하여 복잡하고 급격히 진화하는 데이터 처리환경에 효율적으로 대처할 필요가 있다.

주제어 : 가명정보, 가명처리, 개인정보, 정보주체의 동의, 개인정보자기결정권

• 투고일 : 2022.10.07. / 심사일 : 2022.10.25. / 게재확정일 : 2022.10.26.

Ⅰ. 머리글

가명정보는 그 자체로는 특정개인을 식별할 수 없도록 일정한 변화를 가한 정보이다. 특정개인이 누구인지를 알아볼 수 없도록 특정개인에 관한 정보 중에서 성별, 연령, 상세주소 등의 내용을 삭제하고 실명 대신에 기호나 번호로 대체 및 변경하는 조치가 가명처리이고 이렇게 해서 생성된 정보가 가명정보이다. 개인정보의 식별가능성을 판단하는 기준은 구별가능성(Single-out), 연결가능성(Linkability), 추론가능성(Inference) 등의 3가지 요소이다. 구별가능성(Single-out)은 보유 중인 개인정보 항목에서 특정 개인 1인만을 별도로 분리할 수 있는지 여부를 지칭하고, 연결가능성(Linkability)은 하나의 개인 또는 동일 속성을 공유하는 집단에 관하여 2개 이상의 데이터를 연결할 수 있는지 여부를 의미하고, 추론가능성(Inference)은 2개 이상의 정보가 서로 정확하게 연결되어 있지 않더라도 추론에 의하여 연결이 가능한지 여부를 의미한다. 가명처리는 구별가능성을 제외한 연결가능성과 추론가능성을 제거한 것이다. 익명처리는 개인정보에 대하여 구별가능성, 연결가능성, 추론가능성 등 전기 3요소 모두를 제거하여 특정개인과 연결성을 복원할 수 없도록 한 것이다.[1)]

종래에 가명정보에 관한 논의는 여러 측면에서 익명정보에 관한 논의와 평행하여 전개되어 왔다. 그러나 기술발전에 따라서 완전한 익명처리가 사실상 불가능한 경우가 많고, 어느 한 시점에 익명처리를 하였다고 판단하였다고 하더라도 이후 상황의 진전에 따라 재식별 위험이 커져 다시 개인정보가 될 수 있음이 인식되게 되었다. 익명처리 내지 비식별처리 중심으로 이루어지던 개인정보 보호와 개인정보 이용 사이의 조화에 관한 논의가 가명정보 활용으로 옮

1) 김창조, “정보주체의 개인정보자기결정권 보장과 개인정보의 활용”, 법학논고 75권 2021, 53면 이하.

겨가고 있다. 이러한 것을 반영하여 개인정보보호법은 가명정보에 대하여 보다 적극적으로 규정하게 되었다, 가명정보는 그 자체로는 특정인을 식별할 수 없으므로 그 활용에 대하여는 폭 넓은 길을 열어주고 있다. 가명정보에 대하여는 통계작성, 과학적 연구, 공익적 기록보존 등으로 처리하는 경우에 정보주체의 동의를 받지 않아도 처리할 수 있도록 하였다. 가명정보와 익명정보의 차이가 양자의 상대적 구별을 전제로 하여 전체적으로 비례의 원칙이 적용된다는 점에 비추어 볼 때, 익명정보와 연속선상의 개념으로도 볼 수 있다. 가명정보와 익명정보의 경계가 반드시 뚜렷한 것이 아니고, 다분히 유형적 개념으로 파악될 수도 있다. 그러나 익명정보는 개인정보에 해당되지 않지만, 가명정보는 개인식별성이 제한되지만, 여전히 개인정보보호법의 적용을 받는 개인정보에 해당한다. 이하에서는 개인정보보호법을 중심으로 가명정보에 관한 법제를 분석하고 제도적 발전방향을 검토하려 한다.

Ⅱ. 가명정보개념과 기능

1. 가명정보의 개념

가명정보는 개인을 알아볼 수 있는 정보나 다른 정보와 쉽게 결합하여 알아볼 수 있는 개인정보를 동 정보의 일부를 삭제하거나 일부 또는 전부를 대체하는 등의 방법으로 추가 정보가 없이는 특정 개인을 알아볼 수 없도록 처리하는, 즉 가명처리를 함으로써 원래의 상태로 복원하기 위한 추가정보의 사용ㆍ결합 없이는 특정개인을 알아볼 수 없게 된 정보를 말한다.[2)]

개인정보는 살아있는 개인에 관한 정보로서 세가지 유형으로 규정된다. 첫째 유형은 「성명, 주민등록번호 및 영상 등을 통하여 개인을 알아볼 수 있는 정보」이다. 둘째 유형은 「해당 정보만으로는 특정 개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 알아볼 수 있는 정보」이다. 셋째 유형은 개인정보의 일부를 삭제하거나 일부 또는 전부를 대체하는 등의 방법으로 추가 정보가

2) 개인정보보호법 제2조 1호 다목, 「신용정보의 이용 및 보호에 관한 법률」 제2조 제15호 "가명처리"란 추가정보를 사용하지 아니하고는 특정 개인인 신용정보주체를 알아볼 수 없도록 개인신용정보를 처리(그 처리 결과가 다음 각 목의 어느 하나에 해당하는 경우로서 제40조의2제1항 및 제2항에 따라 그 추가정보를 분리하여 보관하는 등 특정 개인인 신용정보주체를 알아볼 수 없도록 개인신용정보를 처리한 경우를 포함한다)하는 것을 말한다.

없이는 특정 개인을 알아볼 수 없도록 처리한 정보이다.[3] 개인정보의 첫째 유형이 엄격한 요건을 요구하는 반면에, 둘째와 셋째 유형은 느슨한 완화된 요건을 요구하는 것으로 볼 수 있다. 가명정보는 셋째 유형의 정보로서 첫째 유형과 둘째 유형의 정보를 가명처리하여, 즉 개인정보의 식별가능성 중 구별가능성을 제외한 연결가능성과 추론가능성을 배제하여 추가정보 없이는 개인식별성이 제한된 개인정보를 지칭한다.

이 세가지 유형의 개인정보가 아니면 개인정보보호법 상의 관련규제의 적용을 받지 아니한다.

따라서 시간·비용·기술 등을 합리적으로 고려할 때 다른 정보를 사용하여도 더 이상 개인을 알아볼 수 없는 익명정보는 개인정보 보호에 관한 관련규정의 적용을 배제 받게 된다. 개인정보처리자는 개인정보를 익명 또는 가명으로 처리하여 개인정보 수집목적을 달성할 수 있는 경우, 익명처리가 가능한 경우에는 익명에 의하여 처리하고, 익명처리로 목적을 달성할 수 없는 경우에는 가명에 의하여 처리될 수 있도록 하여야 한다.[4]

2. 가명정보제도의 연혁

2020년 개정 이전 개인정보보호법에는 가명정보라는 용어는 사용되지 않았지만 관련 규정을 두고 있었다. 2020년 개정 이전의 개인정보보호법 제18조 제2항 제4호는 「통계작성 및 학술연구 등의 목적을 위하여 필요한 경우로서 특정 개인을 알아볼 수 없는 형태로 개인정보를 제공하는 경우」에는 개인정보의 추가처리가 가능하다고 규정하고 있었다. 여기에서 특정 개인을 알아볼 수 없는 형태로 개인정보를 제공하는 경우라 함은 개인정보에 대하여 가명처리된 개인정보를 제공하는 것으로 해석되었다. 한편 익명처리된 정보는 개인정보에 해당하지 않기 때문에 동법에 따른 개인정보보호규정의 적용을 피할 수 있었다.[5] 동법 제3조 제7항은 「개인정보처리자는 개인정보의 익명처리가 가능한

3) 개인정보보호법 제2조 제1호

4) 개인정보보호법 제3조 제7항

5) 같은 취지의 규정은 2020년 개정 전 「신용정보의 이용 및 보호에 관한 법률」에도 있었다. 동법 제32조 제6항 제4호는 「채권추심(추심채권을 추심하는 경우만 해당한다), 인가·허가의 목적, 기업의 신용도 판단, 유가증권의 양수 등 대통령령으로 정하는 목적으로 사용하는 자에게 제공하는 경우」를 들고 제33조 제1호는 이 경우에 개인신용정보를 동이 없이 이용할 수 있다고 규정하고 있는데, 동법 시행령 제28조 제10항 제7호는 「통계작성 및 학술연구 등을 위하여 필요한 경우로서 신용정보회사 등으로부터 특정 개인을 알아볼 수 없는 형태로 개인신용정보를 제공받기 위한 목적」을 들고 있었다.

경우에는 익명에 의하여 처리될 수 있도록 하여야 한다」고 규정하고 있는데, 여기에서 규정하는 익명의 경우에는 가명처리와 익명처리 기술을 포함하는 것이라고 해석되었다.[6]

개정 개인정보보호법은 가명처리가 허용되는 목적으로 기존의 통계작성은 존치하고, 학술연구가 과학적 연구로 보다 넓은 개념으로 대체되었으며, 공익적 기록보존이 추가되었다. 이러한 점을 고려한다면 개정 개인정보보호법에서는 정보주체의 동의 없이 가명처리가 허용되는 특별한 목적이 보다 확대되었다고 평가할 수 있다. 이는 과학적 연구와 통계적 목적을 위하여 사회의 지식 증가에 대한 합리적 기대가 고려되었기 때문이다. 개인정보처리자가 통계작성, 과학적 연구, 공익적 기록보존 등을 위하여 정보주체의 동의 없이 가명정보를 처리할 수 있게 규정함으로써 적어도 이들 세 가지 목적을 위하여 개인정보보호법 제15조 및 제17조의 법적 근거가 필요로 하지 않게 되었다. 물론 가명처리 되기 이전의 개인정보의 수집에는 개인정보의 수집 · 이용을 규정한 개인정보보호법 제 15조의 법적 근거를 필요로 한다.

가명정보에 대한 법적 규율에 대한 기본적 사항을 제외하고 가명처리를 위한 기술적 사항에 대해서는 가이드라인이라는 형태의 행정지도를 통하여 규율하여 왔다. 가명정보 혹은 데이터 익명화와 관련하여 우리나라 정부기관에서 처음으로 언급한 자료는 2013년 6월 「정부 3.0추진 기본계획」이다. 이 문서에서 공공정보 개발 및 공유에 따른 개인정보보호를 위한 「개인 비식별화(익명화)」처리 기법 보급이라는 표현을 사용한 것이 처음으로 알려져 있다. 그 이후 방통위, 미래부 및 행자부 등의 관련 기관들이 「빅데이터 개인정보보호 가이드라인」, 「빅데이터 활용을 위한 개인정보 비식별화 사례집」, 「개인정보 비식별화 기술활용 안내서」등을 각각 발표하였다. 2016년 7월 정부부처 합동으로 전기 비식별화조치에 관한 기존의 모든 것을 대체하는 「개인정보 비식별화 조치 가이드라인」을 공표하였다.[7] 이 규정은 2020년 개정 개인정보보호법이 시행되기 이전까지 존속하였으나, 개정법의 시행과 함께 공표된 개인정보보호위원회의 「가명정보처리 가이드라인」, 보건복지부와 개인정보보호위원회 공동으로 제정 발표한 「보건의료 데이터 활용 가이드라인」 등에 의해서 대체되게 되었다.

6) 이동진, "가명정보의 특례", 「데이터와 법」(서울: 박영사, 2021), 155-157면.
7) 이기혁 · 김원 · 홍준호, 「개인정보보호와 활용개론」(서울: 생능출판, 2017), 5면.

3. 개인정보자기결정권과 가명정보

개인정보자기결정권은 자신에 관한 정보가 언제 누구에게 어느 범위까지 알려지고 또 이용되도록 할 것인지를 정보주체가 스스로 결정할 수 있는 권리이다. 이 권리에 관하여 1983년 독일연방헌법재판소가 인구조사법(Volkszählungsgesetz) 관련 결정에서 정보의 자기결정권(Recht auf informationelle Selbstbestimmung)을 인정하였다.[8] 그 후 우리 헌법재판소도 개인정보자기결정권을 헌법상 기본권으로 인정하게 되었다.[9]

개인정보자기결정권은 인간의 존엄과 가치, 행복추구권을 규정한 헌법 제10조 제1문에서 도출되는 일반적 인격권 및 헌법 제17조의 사생활의 비밀과 자유에 의하여 보장되는 독립한 기본권이다. 개인정보자기결정권의 헌법적 근거에 대하여 헌법재판소는 다음과 같이 설시하고 있다. 즉, 「인간의 존엄과 가치, 행복추구권을 규정한 헌법 제10조 제1문에서 도출되는 일반적 인격권 및 헌법 제17조의 사생활의 비밀과 자유에 의하여 보장되는 개인정보자기결정권은 자신에 관한 정보가 언제 누구에게 어느 범위까지 알려지고 또 이용되도록 할 것인지를 그 정보주체가 스스로 결정할 수 있는 권리이다. 즉 정보주체가 개인정보의 공개와 이용에 관하여 스스로 결정할 권리를 말한다. 개인정보자기결정권의 보호대상이 되는 개인정보는 개인의 신체, 신념, 사회적 지위, 신분 등과 같이 개인의 인격주체성을 특징짓는 사항으로서 그 개인의 동일성을 식별할 수 있게 하는 일체의 정보라고 할 수 있고, 반드시 개인의 내밀한 영역이나 사사(私事)의 영역에 속하는 정보에 국한되지 않고 공적 생활에서 형성되었거나 이미 공개된 개인정보까지 포함한다」고 하였다.[10] 주관적 공권으로서 개인정보자기결정권은 정보주체가 타인이 보유하고 있는 자신의 정보에 접근하여 열람하거나 그 정보의 정정, 삭제, 차단 등을 요구함으로써 자신에 관한 정보에 통제할 수 있는 권리로서 이를 보장하기 위하여 개인정보보호법은 정보주체에게 동의권, 개인정보의 열람청구권, 개인정보 정정청구권과 삭제ㆍ차단청구권, 개인정보의 처리정지 요구권, 손해배상청구권 등을 인정하고 있다.[11]

이러한 개인정보자기결정권를 배경으로 개인정보보호법에서는 정보주체의 동의를 개인정보처리의 원칙적인 요건으로 정하고 있는데 일반적인 공익보다

---

8) BVerfGE 65, 1.
9) 헌재 2005. 5. 26. 99헌마513 등, 판례집 17-1, 668 [전원재판부]
10) 2005. 7. 21. 2003헌마282・425(병합) 전원재판부
11) 김창조, 전게논문, 54면.

개인정보자기결정권을 우선시키는 것으로 평가할 수 있다. 그러나 동의요건 준수를 위하여 많은 사회적 비용이 발생한다. 이러한 점을 고려한다면 동의 절차를 통하여 정보주체가 합리적인 선택을 하는지 여부 및 정보주체의 동의를 요구하는 것이 개인정보자기결정권을 보호하는 가장 적절한 방법인지 여부에 대해서는 의문이 제기될 수도 있다. 2020년 개정 개인정보보호법은 주관적 공권으로서 개인정보자기결정권을 보장하면서 법적 소여 변화에 따른 개인정보 활용의 확대가 가능하도록 동의요건 완화를 도모하고 있다. 이를 위하여 개정 전 개인정보보호법에서는 가명에 대한 명확한 규정을 두지 않았으나 새로운 개인정보보호법에서는 가명정보의 정의규정을 추가하고 개인정보처리자가 통계작성, 과학적 연구, 공익적 기록보존을 위하여 정보주체의 동의없이 가명정보를 처리할 수 있게 하였다.

4. 가명정보제도의 기능

가명처리는 정보주체의 프라이버시를 보호하는 수단이 된다. 가명처리가 데이터 활용의 수단으로만 볼 수도 있지만, 가장 중요한 기능은 정보주체의 프라이버시를 보호하면서 데이터를 안정되게 활용하기 위한 것이다. 예컨대 어떤 기관 내의 IT 개발부서가 소프트웨어를 개발·테스트하는 상황을 생각해 보자. 개발부서가 정보주체에 대한 직접적 정보를 보유하고 있을 필요가 없는 경우도 많다. 그러나 실제 상황과 전혀 동떨어진 데이터를 가지고는 제대로 된 개발이나 테스트를 진행하기 어려울 경우도 적지 않을 것이다. 이때 좋은 실무관행은 가명처리를 하여 실제 개인정보와 유사한 외견을 갖추고 있지만, 개인을 식별할 수 없는 형태로 변형하여 이를 활용하는 것이다. 이렇게 함으로써 혹시 IT 개발부서에서 개인정보가 유출되더라도 프라이버시 침해의 위험성을 줄어들게 할 수 있다. 또한 IT 개발부서 개발자가 임의로 특정 개인정보를 검색하려고 시도하는 경우를 막을 수도 있다. 이러한 의미에서 가명처리를 하여 데이터를 활용하는 것이 바람직하다고 할 수 있다.

가명처리를 활용하는 또 다른 중요한 기능은 정보주체의 동의를 얻지 않고 당해 정보를 활용하거나 제3자에게 제공하는 것이다. 만약 정보주체의 동의를 얻는 것이 어렵지 않다면 정보주체의 동의를 얻어 사용하거나 제3자에게 제공하는 것이 우선적으로 고려하여야 할 것이다. 장래에 수집되는 데이터에 대해서는 새로운 동의를 얻는 것이 어렵지 않을 수도 있다. 동의를 받는 방법을 변경하면 가능하기 때문이다. 하지만 이미 수집해 둔 데이터에 대해서는 일일

이 정보주체로부터 다시 동의를 얻는 것이 매우 어려울 수 있다. 이러한 때에는 기존 데이터를 가명처리하여 가명정보의 형태로 활용할 것을 고려하게 될 것이다.[12]

5. 가명정보와 익명정보

개인정보보호법 제58조의2는 익명정보에 대해 「이 법은 시간·비용·기술 등을 합리적으로 고려할 때 다른 정보를 사용하여도 더 이상 개인을 알아볼 수 없는 정보에는 적용하지 아니한다」고 규정한다.[13] 익명정보는 개인식별가능성이 없는 정보로서 더 이상 개인정보에 해당하지 않는 정보이다.[14] 익명정보가 되기 위해서는 개인정보처리자가 잠재적으로 정보를 식별할 가능성이 있는 제3자가 추가정보를 입수할 가능성 등 개인을 알아보는 데 소요되는 시간, 비용, 기술 등을 합리적으로 고려할 때 손쉽게 식별할 수 없는 정도에 이르러야 한다. 익명처리가 되어 익명정보가 되면 개인정보보호법의 적용대상이 되지 않는데 비하여 가명정보는 개인정보로서 개인정보보호법의 적용대상이 된다. 전술한 바와 같이 개인정보처리자는 개인정보를 익명 또는 가명으로 처리하여도 개인정보 수집목적을 달성할 수 있는 경우 익명처리가 가능한 경우에는 익명에 의하여, 익명처리로 목적을 달성할 수 없는 경우에는 가명에 의하여 처리될 수 있도록 하여야 한다.

가명정보와 익명정보는 모두 개인식별가능성을 제한하였다는 점에서 공통적이다. 그러나 양자의 활용목적에서 차이가 있다. 가명정보는 그 활용목적이 통계작성, 과학적 연구, 공익적 기록보존 등으로 제한되어 있다. 가명정보는 이러한 목적을 수행하기 위한 이들에게 한정적으로 제공되고, 그 목적으로만 사

12) 김병필, “가명처리와 가명정보의 이해”, 「보건의료와 개인정보」(서울: 박영사, 2021), 84-85면.

13) 「신용정보의 이용 및 보호에 관한 법률」은 제2조 제17호에서 더 이상 특정 개인인 신용정보주체를 알아볼 수 없도록 개인신용정보를 처리하는 것을 익명처리라고 정의하고 익명처리된 정보를 익명정보라고 한다.

14) 「신용정보의 이용 및 보호에 관한 법률」에 따르면 익명처리는 「더 이상 특정 개인인 신용정보주체를 알아볼 수 없도록 개인신용정보를 처리하는 것」을 의미하고(동법 제2조 제17호), 신용정보회사 등은 개인신용정보에 대한 익명처리가 적정하게 이루어졌는지 여부에 대하여 금융위원회에 그 심사를 요청할 수 있다(동법 제40조의2 제3항). 그리고 금융위원회가 제3항의 요청에 따라 심사하여 적정하게 익명처리가 이루어졌다고 인정한 경우 더 이상 해당 개인인 신용정보주체를 알아볼 수 없는 정보로 추정한다고 규정하고 있다(동법 제40조의2 제4항).

용된다. 주로 조직 내부에서 통계처리나 데이터 분석을 담당하는 다른 부서에 제공하거나 다른 연구자·연구기관에 제공될 것을 염두에 두고 있다. 이에 비하여 익명정보는 개인식별가능성이 없는 정보를 의미하므로 그 활용목적에 제한이 없다. 익명정보는 다양한 정보활용가능성이 존재하고 개인정보보호법의 적용도 배제되므로 누군가 익명정보로부터 개인을 식별하고자 하는 상황까지 충분히 대비하여 개인정보를 익명정보로 변형하기 위해서는 더욱 강력하고 엄격한 비식별조치가 취해져야 한다.

개인정보처리자가 정보를 목적 외로 이용하는 경우에 추가정보를 모두 삭제하여야 비로소 익명정보가 되는 것이 일반적이다. 원본 정보가 존재하는 한 익명정보와 원본정보의 대조를 통하여 누구에게 귀속되는 정보인지 식별가능성이 높기 때문이다. 그렇지 않다면 원칙적으로 가명정보가 된다. 그러나 제3자에게 제공하는 경우, 추가정보를 가지고 있는 제공하는 측에서 보면 가명정보인 반면에, 추가정보가 없고 추가정보가 적절히 관리되고 있어 쉽게 입수할 수도 없이 제공받는 측으로서는 익명정보일 수 있다. 이러한 경우에 가명정보로 취급하는 것은 항상 합리적이라고 할 수 없다. 경우에 따라서는 제공받는 입장에서 익명정보로 보아서 통계작성 등의 목적에 구애받을 필요가 없는 경우도 있을 수 있다. 일본의 2020년 개정법은 그러한 취지를 명문으로 규정하였다.[15]

여기에서 문제되는 것이 추가정보의 의미이다. 당해 정보만으로 정보주체가 누구인지를 알아볼 수 없기만 하면, 가명정보가 되는 것인가 그렇지 않으면 손쉽게 입수할 수 있는 다른 추가정보를 고려해야 하는가라는 점이다. 어떤 정보로부터 그 정보주체를 추론하는 데 쓸 수 있는 추가정보 중에는 개인정보처리자가 분리 보관하는 것 외에도 시중에서 어렵지 않게 입수할 수 있는 것이 있을 수 있다. 가명정보가 되기 위해서는 이들까지 고려하여 모든 식별가능성을 없애야 한다는 해석을 할 수도 있다. 즉 개인정보처리자가 분리 보관한 추가정보 이외의 추가정보도 모두 고려하여 가명처리의 실행 여부를 판단하여야 한다는 것이다. 이러한 접근방법을 취할 경우, 식별자를 모두 제거하였어도 다른 정보를 통하여 정보주체를 식별할 수 있을 때에는 가명처리가 되었다고 할 수 없고 더 많은 처리를 하여야 한다. 가명정보 인정과정에서 추가정보의 범위를 과도하게 확장하여 해석할 경우, 가명정보 인정 범위가 극단적으로 제한되어 가명정보를 인정하는 제도적 의미가 반감할 수도 있다. 가명제도를 인정하여 도입한 취지.목적을 고려하여 추가정보의 인정 범위를 적정하게 획정할 필요가

15) 이동진, 전게논문, 155면.

있다.[16]

Ⅲ. 가명정보 특례의 적용범위

2020년 개정 개인정보보호법은 제3절 가명정보의 특례라는 표제 하에 제28조의2 내지 제28조의7의 일련의 규정을 신설하였다. 가명정보 특례가 중요한 까닭은 이 규정이 사적 섹터에서 개인정보를 목적 외에 이용 및 제공할 수 있는 거의 유일한 예외이기 때문이다. 개인정보보호법 제18조 제2항 각호는 목적 외 이용 및 제공사유를 열거하고 있다. 그중에서 동의(제1호)와 급박한 생명, 신체, 재산상의 이익의 보호(제3호)를 제외한 법률규정(제2호), 법령준수(제5호), 조약, 그 밖의 국제협정의 이행(제6호), 범죄의 수사와 공소의 제기 및 유지(제7호), 재판(제8호), 형의 집행(제9호) 등은 모두 공적 섹터에 관련된 것이다. 개인정보보호법은 가명정보 특례에 관하여 개인정보처리자는 통계작성, 과학적 연구, 공익적 기록보존 등을 위하여 정보주체의 동의 없이 가명정보를 처리할 수 있다고 규정하고 있다. 이 규정은 공적섹터에서도 활용이 기대되지만, 그보다 그 광범위한 영역의 사적 섹터에서 활용이 기대되고, 오히려 사적 섹터에서 활용이 적합한 규정이라고 할 수 있다.[17]

1. 통계작성

통계작성은 특정집단이나 대상 등에 관하여 작성한 수량적인 정보를 의미한다. 통계작성을 위한 가명처리란 통계를 작성하기 위해 가명정보를 이용, 분석, 제공하는 등의 가명처리를 지칭한다고 한다. 개인정보처리자가 정보주체의 동의 없이, 통계처리를 위하여 가명정보를 처리하는 경우, 그 목적은 반드시 공익적 통계작성에 국한되지 않고 상업적 목적의 통계작성도 포함한다.[18]

한편 통계법에 따르면 통계작성이라 함은 「통계작성기관이 정부정책의 수립·평가 또는 경제·사회현상의 연구·분석 등에 활용할 목적으로 산업·물가·인구·주택·문화·환경 등 특정의 집단이나 대상 등에 관하여 직접 또는 다른 기관이나 법인 또는 단체 등에 위임·위탁하여 작성하는 수량적 정보를

16) 이동진, 전게논문, 163면.
17) 이동진, 전게논문, 150면.
18) 개인정보보호위원회, 가명정보처리 가이드라인, 2022. 4, 11면.

말한다」[19]고 규정하고 있다. 통계의 작성과정에서 알려진 사항으로서 개인이나 법인 또는 단체 등의 비밀에 속하는 사항은 보호되어야 하며, 통계의 작성을 위하여 수집된 개인이나 법인 또는 단체 등의 비밀에 속하는 자료는 통계작성 외의 목적으로 사용되어서는 아니 된다.[20] 해당 통계자료를 다른 자료와 대응 또는 연계함으로써 특정의 개인이나 법인 또는 단체 등의 식별이 가능하게 되는 경우에는 통계기관의 장은 통계자료를 제공하지 아니할 수 있다고 한다.[21]

2. 과학적 연구범위

가명정보를 정보주체의 동의 없이 활용할 수 있는 것은 통계작성, 과학적 연구, 공익적 기록보존 등의 목적을 달성하기 위한 경우로 제한된다. 이러한 목적 범위에 어디까지 포함되는 것인지가 쟁점이 된 것이 과학적 연구의 범위이다. 2020년 개인정보보호법 개정과정에서도 가명정보의 활용 목적범위에 관한 논란이 되었고, 개정 후에도 논란이 계속되고 있다. 가명정보의 활용범위를 좁게 보아야 한다는 입장에서는 과학적 연구의 범위에서 상업적·산업적 연구를 제외시켜야 한다는 주장을 제기하였다. 이러한 입장에서는 가명정보는 정보주체의 동의 없이 활용하는 것이니 공동체 전체에 이익이 귀속될 수 있는 경우에 한정해야지, 특정 기업의 이익을 위해서 활용하도록 허용해서는 안 된다고 한다.[22]

개인정보보호법 제2조 제8호 과학적 연구를 「기술의 개발과 실증, 기초연구, 응용연구 및 민간 투자 연구 등 과학적 방법을 적용하는 연구」라고 정의하고 있다. 과학적 연구란 연구방법에 관한 것이고, 상업적 연구는 연구결과를 영리목적으로 활용하는 경우를 의미하는 것이므로 다른 층위의 개념이다. 따라서 현행법 해석상 상업적 연구라고 해서 과학적 연구에 포함되지 않는다고 해석하기 어렵다. 더욱이 현실적으로 대학, 공공연구기관과 기업들 간에 긴밀한 협업 연구가 이루어지고 있는 상황에서 어떤 연구가 상업적인지 아니지 나누기도 어렵다. 따라서 현행 개인보호보호법상으로 과학적 연구에 기업에 의한 상업적 연구를 제외한다고 해석하기는 쉽지 않다고 볼 수 있다. 가명처리가이드

19) 통계법 제3조 제1호
20) 통계법 제33조 제2항
21) 통계법 제31조 제3항
22) 김병필, 전게논문, 85면

라인에서도 과학적 연구와 관련하여 공적 자금으로 수행하는 연구뿐만 아니라 민간으로부터 투자를 받아 수행하는 연구에서도 가명처리가 가능하다고 한다.[23]

3. 공익적 기록보존

개인정보처리자는 공익적 기록보존을 위하여 정보주체의 동의 없이 가명정보를 처리할 수 있다. 공익적 기록보존이란 공공의 이익을 위하여 지속적으로 열람할 가치가 있는 정보를 기록하여 보존하는 것을 의미한다.[24] 공익적 기록보존 목적으로 가명정보가 처리되는 경우는 모든 기록이 아니라 기록보존이 법으로 요구된 경우에 한정될 것이다. 또한 공익적 기록보존은 역사적 의미를 가지는 점에서 사망한 사람에 관한 정보인 경우가 많겠지만, 가명정보 규정의 적용을 받는 개인정보는 살아있는 개인에 관한 정보이다. 이 점에서 공익적 기록보존을 목적으로 처리되는 가명정보는 사망한 사람에 관한 정보는 포함되지 않을 것이다. 다만 해당 기록이 사망한 사람과 살아 있는 사람에 관한 정보를 포함하는 경우에는 살아 있는 사람에 관한 정보가 존재하는 점에서 공익적 기록보전 목적의 가명정보처리가 허용될 수 있을 것이다.[25]

Ⅳ. 가명처리

1. 가명정보와 가명처리

가명처리란 개인정보의 일부를 삭제하거나 일부 또는 전부를 대체하는 등의 방법으로 추가 정보 없이는 특정 개인을 알아볼 수 없도록 처리하는 것을 말한다.[26] 가명정보는 당해 정보만 보아서는 그 정보가 귀속되어 살아있는 개인이 누구인지 알 수 없어야 한다. 이름이나 주민등록번호, 학번, 사진 등만으로 또는 그들 중 몇이 함께 제시되는 경우에 개인정보처리자가 정보주체가 누구인지 알아볼 수 있다면 가명처리가 되었다고 할 수 없다. 그러한 정보 중 일

23) 개인정보보호위원회, 가명정보처리 가이드라인, 2022. 4, 12면.
24) 개인정보보호위원회, 가명정보처리 가이드라인, 2022. 4, 13면.
25) 박노형, 「개인정보보호법」(서울: 박영사, 2021), 246면.
26) 개인정보보호법 제2조 제1의2호

부를 삭제하거나 대체함으로써 더는 정보주체를 곧바로 알 수 없게 하여야 가명처리가 된 것이다.

가명처리에는 본래의 값을 랜덤 값으로 대체하는 방법과 해시함수나 암호화를 통한 방법 등이 있다. 가장 많이 사용하는 가명처리 기법으로 비밀키가 있는 암호화, 해시함수, 보관키가 있는 해시함수 및 결정성 암호화나 키를 삭제한 키해시함수, 토큰화를 들 수 있다. 비밀키가 있는 암호화의 경우 가명정보 집합에 개인정보가 암호화된 형태로 포함되어 있어 키 보유자가 어렵지 않게 재식별을 할 수 있다. 그 반면에 해시함수의 경우 입력값의 범위를 알고 있는 때에는 복구가 어렵지만 가능하고, 보관키가 있는 해시함수의 경우 보관키를 보유한 사람은 쉽게 개인정보를 복구할 수 있으나 보관키가 없으면 복구가 훨씬 어렵다고 한다. 결정성 암호화 내지 키를 삭제한 키해시함수는 랜덤 대체에 상응하여 현재의 계산능력으로는 재식별하기 어렵고, 토큰화는 금융업에서 흔히 쓰는 기법으로 카드 ID를 다른 숫자로 대체하는 것인데 일반적으로 일방향 암호화, 일련번호 또는 랜덤값을 쓴다고 한다.[27]

신용정보의 이용 및 보호에 관한 법률」 제2조 제15호는 가명처리란 추가정보를 사용하지 아니하고는 특정 개인인 신용정보주체를 알아볼 수 없도록 개인신용정보를 처리하는 것이라고 하면서 그 처리 결과가 「어떤 신용정보주체와 다른 신용정보주체가 구별되는 경우」 또는 「하나의 정보집합물에서나 서로 다른 둘 이상의 정보집합물간에서 어떤 신용정보주체에 관한 둘 이상의 정보가 연계되거나 연동되는 경우」 등에 해당하는 경우로서 신용보호법 제40조의2 제1항 및 제2항에 따라 그 추가정보를 분리하여 보관하는 등 특정 개인인 신용정보주체를 알아볼 수 없도록 개인신용정보를 처리한 경우를 포함한다고 한다. 하나의 정보집합물에서 또는 연계된 복수의 정보집합물에서 하나 또는 복수의 정보가 A(일련번호 0001)라는 정보주체에 귀속되고 다른 정보가 B(일련번호0002)라는 정보주체에 귀속되는 것을 알 수 있더라도 실제로 A, B((일련번호 0001, 0002)가 누구인지 알 수 없다면 가명정보가 된다. 이 경우 A, B가 누구인지를 알 수 있게 해주는 추가정보는 분리하여 보관하거나 삭제하여야 한다. 이러한 추가정보에는 다양한 정보가 포함될 수 있다. 전형적으로는 매칭 테이블(matching table)은 물론이고, 해시함수가 사용된 경우나 암호화된 경우 키(key)도 추가정보에 해당한다. 그러므로 당해 정보만 보아서는 정보주체가 누구인지 알 수 없다 하여도 그것을 알 수 있는 추가정보가 같은 개인정보처

27) 이동진, 전게논문, 152면.

리자에게 있다면 내부적으로 추가정보의 통제하는 장치(패스워드, 망분리 등)를 마련하여야 비로소 가명정보가 된다.[28]

2. 가명처리할 수 있는 대상정보 범위

가명정보의 특례의 적용대상이 될 개인정보는 통상의 개인정보 이외에 민감정보나 고유식별정보가 포함된 정보도 적용대상이 될 수 있다. 이 점이 2020년 개정의 가장 큰 변화 중의 하나라고 할 수 있다. 2020년 개정 전 개인정보보호법 제18조 제2항 제4호는 제3장 제1절에 규정되어 있어 제2절의 민감정보와 고유식별정보에 대하여 적용되는지 여부가 문제될 수 있었는데 개정법은 이를 삭제한 대신 제3절에 가명정보의 특례를 넣어 체계적으로 민감정보와 고유식별정보에 적용됨을 분명히 한 것이다. 이것은 비교법적으로도 타당하고 내용적으로도 합리적 개정이라고 평가된다.[29] 가령 건강정보는 광범위하게 민감정보로 분류되고 있고, 정치적 사상 등에 관한 정보도 그러한데 이러한 것이 가명정보로도 통계작성, 과학적 연구 또는 공익적 기록보존의 대상이 되지 않고 오직 동의 없이는 사용이 제한된다고 해석하는 것은 합리성을 결한다고 보여 진다.

개인정보보호법과 신용정보보호법은 가명처리를 의무화한 점에 특징이 있으므로 가명처리하지 않으면, 동의 없이 통계작성 등의 목적만으로 개인정보를 이용할 수 없다. 최소한의 가명처리 만으로 동의 없는 가명정보의 추가처리가 당연히 적법해지는 것은 아니다. 개인정보보호법 제3조 제7항 등에 반영되어 있는 비례원칙과 데이터 최소화 등이 여기에도 고려되어야 한다. 개인정보처리자는 목적을 달성할 수 있는 한 불필요한 개인정보노출을 최소화하여야 하고 더 높은 수준의 가명처리를 할 의무를 진다. 개별적 사안에서 어느 정도의 가명처리가 적절한지의 판단은 동태적·맥락 의존적이므로 실체적 기준만으로 접근하기보다는 절차적 기준을 결합 내지 보완하는 것이 바람직하다.[30] 가명정보의 결합에 관한 개인정보보호법 제28조의3도 결합기관의 장의 승인을 거치게 함으로써 일정한 절차적 요소를 예정하고 있다. 개인정보처리자는 가명정보를 제3자에게 제공할 때 특정개인을 알아보기 위하여 사용될 수 있는 정보, 즉 추가정보를 제공하여서는 안 된다.[31]

---

28) 이동진, 전게논문, 161-162면.
29) 이동진, 전게논문, 167면.
30) 이동진, 전게논문, 165-168면.
31) 개인정보보호법 제28조의2 제2항

3. 가명처리 과정과 가명처리방법

1) 가명처리과정

개인정보처리자는 개인정보 처리 기본원칙을 준수하는 범위 내에서 가명처리의 전 과정을 진행해야 한다. 개인정보처리자는 가명정보 처리목적과 처리환경을 고려하여 가명처리방법을 자체적으로 판단할 수 있다. 개인을 식별할 가능성이 높은 정보는 삭제하거나 원래의 정보로 복원할 수 없도록 처리하고, 그 외의 정보는 가명정보 처리목적 달성을 위해 적절한 처리방법을 선택하여 가명처리할 수 있다. 개인정보처리자가 가명정보 처리 목적에 필요한 최소한의 정보만을 처리하고, 가명처리 과정에서 재식별 가능성이 없는지를 확인하여야 한다. 통상적으로 가명정보 처리절차는 사전준비단계, 처리대상의 위험성검토, 가명처리, 적정성 검토, 안전한 관리 등 5단계로 구분하여 처리된다.[32]

제1단계(목적 설정 등 사전준비) : 사전준비단계에서는 가명정보 활용목적을 명확화·구체화하고, 필요한 서류를 작성한다. 이때 제3자 제공의 경우에는 계약체결 관련서류를 마련한다,

제2단계(처리대상의 위험성 검토) : 여기에서는 사전준비단계에서 설정된 목적을 달성하기 위해 필요한 항목을 개인정보에서 선정하고, 가명처리 대상 데이터의 식별 위험성을 분석·평가하여 가명처리 방법 및 수준을 반영하기 위한 절차이다.

제3단계(가명처리) : 개인정보처리자는 식별위험성 검토 결과를 기반으로 가명정보의 활용 목적달성에 필요한 가명처리 방법 및 수준을 정하여 항목별 가명처리 계획을 설정하고, 항목별 가명처리계획을 기반으로 가명처리를 수행한다. 가명처리과정에서 생성되는 추가정보는 원칙적으로 파기하고 필요한 경우 가명정보와 분리하여 별도로 저장하여야 한다.

제4단계(적정성 검토) : 1, 2, 3단계의 가명처리가 적정하게 수행되었는지 확인하고, 가명처리 한 결과가 가명정보의 처리목적을 달성하기 위해 적절한지 등을 검토한다. 적정성 검토는 (필요서류)→(처리목적의 적합성)→(식별위험성)→(가명처리 방법 및 수준의 적정성)→(가명처리의 적정성)→(처리목적 달성가능성)의 순서로 단계적으로 진행한다. 적정성 검토 시 위원장을 선정하여 절차에 따라 검토를 진행할 수 있도록 하고, 종합적인 내용과 각 검토위원의 의견을 고려한 최종검토결과 및 종합검토의견을 개인정보처리자에게 제출한다.

32) 개인정보보호위원회, 가명정보처리 가이드라인, 2022. 4, 9-38면.

제5단계(안전한 관리) : 마지막으로 안전한 관리 단계에서는 적정성 검토 이후 생성된 가명정보는 법에 따라 기술적 · 관리적 · 물리적 안전조치 등 사후관리를 이행하여야 한다.

2) 가명처리 방법

(1) 식별자와 속성값

전술한 바와 같이 개인정보보호법상 가명처리의 정의규정에 따르면 가명처리란 개인정보의 일부를 삭제하거나 일부 또는 전부를 대체하는 등의 방법으로 추가정보 없이는 특정개인을 알아볼 수 없도록 처리한 것이다. 이 규정에 따르면 가명처리 방법은 「개인정보의 일부를 삭제하거나 일부 또는 전부를 대체하는 등의 방법」을 지칭한다. 그러나 여기에는 「등」이 포함되어 있으므로 가명처리 방법이 개인정보의 삭제 또는 대체에만 국한되는 것이 아니다. 여기서 쟁점이 되는 것은 삭제 또는 대체 등의 대상이 되는 개인정보의 범위가 어디까지인가 하는 것이다. 다음으로 위 조항은 가명처리한 결과물이 충족해야 할 요건을 정하고 있다. 즉, 가명처리의 결과물이 「추가정보 없이는 특정개인을 알아볼 수 없도록」 해야 한다 그러나 여기에서 추가정보가 무엇인지는 법상 정의되어 있지 않다.

가명처리를 위해서 삭제 또는 대체 등의 대상이 되는 정보의 범위를 정하는데 유용한 방법은 식별자(직접식별자)와 속성값(간접식별자)을 구분하는 것이다. 식별자(identifier)는 단독으로 개인을 알아볼 수 있는 정보이고, 해당 정보만으로 특정개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 알아볼 수 있는 것이 속성값(attribute value)이다.[33]

(2) 식별자(직접식별자)에 대한 가명처리

식별자란 개인을 유일하게 식별할 수 있는 정보이다. 사람에게 이름을 붙이는 것은 개인을 식별하기 위한 것이다. 이름 이외에도 개인을 식별하기 위해 붙여 놓은 번호들은 많다. 여권번호, 외국인등록번호, 운전면허번호, 의료기록번호, 건강보험번호 등이 이것에 해당한다. 그 외에도 전화번호, 상세주소, 전자우편주소 등을 통해서도 개인을 식별해 낼 수 있다. 이러한 값들을 식별자라 부른다. 식별자를 개인정보보호위원회의 가명정보처리 가이드라인은 이를 식별

33) 김병필, 전게논문, 74-75면.

정보라고 칭한다.[34]

식별자에 대한 가명처리 방법은 간명하다. 이를 삭제하거나, 원래 정보를 확인할 수 없는 일련번호로 대체하면 된다. 필요한 경우에는 암호화 기법을 이용하여 원래 정보를 복원할 수 없도록 암호화할 수도 있다. 이러한 일련번호나 암호화된 정보가 가명에 해당한다. 이는 가명처리의 가장 기본적인 방법이고, 가명처리라는 용어의 유래이기도 하다.[35]

(3) 속성값(간접식별자)에 대한 가명처리

데이터에는 식별자 이외에 다수의 속성값이 포함되어 있다. 그런데 속성값 중에는 다른 정보와 조합함으로써 개인을 식별해 낼 수 있는 것도 있지만, 그렇지 않는 것도 있다. 속성값 중에서 다른 정보와 결합하여 개인을 식별할 수 있는 정보를 간접식별자라고 할 수 있다. 간접식별자를 지칭하는 용어는 다양하다. 간접식별자에 해당하는 속성값을 가명정보처리 가이드라인은 식별가능정보라 칭하고 있다.[36]

간접식별자와 관련개념으로 특이정보가 있다. 특이정보란 「전체데이터에 식별가능성을 가지는 고유(희소)값, 편중된 분포를 가지는 단일ㆍ다중이용항목」을 의미한다.[37] 이 정보는 그 정보만으로 개인을 식별할 정보가 아니더라도 고유한 특성 때문에 개인식별가능성이 높은 정보이다. 가명처리과정에서 특이정보에 대하여는 이용목적상 반드시 필요하지 않다면 삭제하고 필요하면 상하단을 코딩처리하거나 범주화 등의 처리를 한다.[38] 구체적으로는 「희귀성씨 등 특이한 값, 국내 최고령 등 극단값, 특정데이터 분석집단에서 희소값 등」의 정보가 이에 해당한다. 특이정보를 파악하기 위해서는 데이터셋을 관찰해야 한다. 즉 데이터의 주된 분포 범위 이외에 존재하는 항목들이나 빈도가 매우 적은 항목을 찾아내야 한다.

실무상으로는 데이터에 포함된 속성값 중 어느 항목까지 간접식별자로 볼 것인가 논란이 된다. 이러한 판단이 어려운 이유는 데이터가 사용되는 맥락에 따라 그 판단이 달라지기 때문이다. 예컨대 직업정보는 간접식별자로 인정될 수도 그렇지 않을 수도 있다. 단순히 공무원, 회사원, 자영업 등으로 구분된 경

34) 개인정보보호위원회, 가명정보처리 가이드라인, 2022. 4, 16면.
35) 김병필, 전게논문, 76면.
36) 개인정보보호위원회, 가명정보처리 가이드라인, 16면.
37) 개인정보보호위원회, 가명정보처리 가이드라인, 16면.
38) 개인정보보호위원회, 가명정보처리 가이드라인, 27면

우라면, 다른 정보와 결합하더라도 개인을 식별하기 쉽지 않을 것이다. 하지만 국회의원 등 특수한 직업이 포함된 경우라면 간접식별자로 인정할 필요가 있을 것이다. 여러 변수를 조합해서 개인을 식별할 가능성도 고려해야 한다. 가령 데이터셋에 각각의 직업마다 대략 300명 정도가 포함되어 있는데, 특정 지역에는 해당 직업 종사자가 1-2인의 극소수만 존재한다면 지역과 직업정보를 조합하여 개인을 재식별할 수 있다.

간접식별자를 가명처리하는 데는 다양한 방법이 활용될 수 있다. 통상적인 방법은 데이터의 정밀도를 낮추어서 재식별 위험을 낮추는 것이다, 이를 보통 일반화 또는 범주화라 한다. 예컨대 소득금액 5500만원을 소득 5000만원-6000만원으로 대체 표기하는 것이다. 문자로 저장된 데이터는 더 상위 개념으로 대체할 수 있다.

하지만 데이터를 일반화하더라도 특이정보에 대해서는 여전히 재식별 위험이 남아 있을 수 있다. 이 경우 전술한 바와 같이 특이정보를 삭제하거나 일반화시켜야 한다. 특이정보를 포함한 기록에 대해서는 해당 특이정보 필드 값만을 삭제할 수도 있고, 해당 특이정보를 포함하고 있는 행 전체를 삭제할 수도 있다.[39]

4. 기술적 의미의 가명처리와 법적 의미의 가명처리

가명처리의 일차적 작업은 직접식별자를 삭제하거나 일련번호로 대체하는 것이다. 하지만 데이터셋에서 직접식별자를 그저 삭제해 버릴 수 없는 상황도 존재한다. 가령 일정한 집단 소속 개인에 대해 매년 정보를 수집한 패널 데이터를 작성하는 경우를 생각해 보자. 패널 데이터를 가명처리하여 연구자에게 제공한 다음, 이듬해에 데이터를 업데이트하려면, 기존 데이터와 새 데이터를 연결시킬 수 있는 방법이 필요하다. 또 다른 예는 가명처리된 데이터를 다른 데이터와 결합하는 경우이다. 예컨대 유아 의료정보와 교육정보를 결합하여 유아건강과 학력 성취도 사이의 상관관계를 분석하고자 하는 상황을 생각해 보자. 이러한 분석을 하려면 두 가지 데이터셋에서 개인을 연결시킬 방법이 있어야 한다. 이상과 같은 상황에서는 데이터셋에서 개인을 유일하게 식별할 수 있는 식별자를 남겨 둘 필요가 있다. 하지만 식별자를 원본 그대로 유지할 수는 없으니 이를 암호화하는 것이 일반적이다. 그리고 식별자를 암호화한 값이 가

39) 김병필, 전게논문, 76-78면.

명이 된다.

가명처리에 활용할 수 있는 암호화 기법에는 일방향 암호화와 쌍방향 암호화의 2가지가 있다. 일방향 암호화는 수학적 기법을 활용하여 원본 값을 복원할 수 없는 다른 일련의 숫자 값으로 변형하는 것이다. 일방향 암호화 방법의 단점은 추후 원본 값으로 복원해야 할 경우를 대비해서 원본 값과 암호화된 값을 대응시키는 표(매핑 테이블, mapping table)를 유지해야 한다는 것이다. 만약 방대한 데이터를 가명처리해야 한다면 매핑 테이블을 유지하는 것도 상당한 부담이 될 수 있다. 또 다른 방법은 쌍방향 암호화이다. 암호 키를 이용하여 암호화함으로써, 암호 키를 이용하면 원본 값을 복원할 수 있는 방법이다. 쌍방향 암호화를 이용하면 매핑 테이블을 유지할 필요 없이 암호 키만 있으면 원본 값을 복원해 낼 수 있다. 그 대신 암호 키를 안전하게 보관하여야 한다는 부담이 발생하게 된다.

암호화를 하면 긴 숫자열이 생성되므로 원본 식별자와는 형태가 달라지게 된다. 그런데 IT 부서에서 소프트웨어 개발목적으로 가명정보를 활용하는 경우, 데이터 포맷이 달라지는 것이 불편할 수 있다. 미리 정해진 포맷을 준수하는 경우에만 데이터베이스에 저장할 수 있도록 프로그램된 경우가 적지 않기 때문이다. 예컨대 신용카드 번호는 15자리 또는 16자리 숫자이므로, 이 길이에 맞추는 경우에만 데이터베이스에 저장할 수 있도록 정해 놓은 것이다. 만약 암호화한 값이 미리 정해 놓은 자리수보다 길어진다면, 기존 데이터베이스의 포맷을 변경해야 한다. 이것은 IT 부서의 개발자가 선호하지 않은 일이다. 이러한 문제를 해결하기 위해서 형태보존 암호화 기법을 사용할 수 있다. 형태 보존 암호화를 이용하면 원본 데이터와 암호화된 데이터가 동일하게 된다. 일방향 암호화이든 쌍방향 암호화이든 가명처리 후 원래 정보로 복원하기 위한 추가적인 정보가 남아 있게 된다. 일방향 암호화에서의 매핑테이블, 쌍방향 암호화에서의 암호키가 여기에 해당한다. 이처럼 원래 정보로 복원하기 위한 추가적인 정보를 흔히 가명처리 비밀이라고 한다. 가명처리 비밀이 유출되면 가명처리로부터 원래 정보를 복원해 낼 수 있으므로 안전하게 유지할 필요가 있다. 이상의 논의와 같이 기술적 의미에서의 가명처리는 직접식별자를 다른 값(가명)으로 대체하는 것만을 의미한다.

그러나 법적 의미의 가명처리는 기술적 의미의 가명처리와 구별할 필요가 있다. 법적 의미의 가명처리는 이러한 기술적 의미의 가명처리 이외에도 직접식별자를 삭제 또는 일련번호로 대체하고 속성 값 중 개인식별성이 높은 정보는 일반화(범주화)하거나 삭제하는 등의 조치를 포함하는 것으로 이해된다.[40]

5. 가명처리 후 안정성 확보를 위한 개인정보처리자의 조치의무와 정보주체의 권리제한

가명정보에서 정보주체를 재식별하는 행위는 금지되어 있다. 가명정보를 처리하는 과정에서 재식별을 가능하게 하는 추가정보가 생성된 경우에는 즉시 처리를 중지하고 지체 없이 이를 회수·파기하여야 한다.[41]

개인정보처리자는 가명정보를 처리하는 경우에는 원래의 상태로 복원하기 위한 추가 정보를 별도로 분리하여 보관·관리하는 등 해당 정보가 분실·도난·유출·위조·변조 또는 훼손되지 않도록 대통령령으로 정하는 바에 따라 안전성 확보에 필요한 기술적·관리적 및 물리적 조치를 하여야 한다. 개인정보처리자는 가명정보 및 가명정보를 원래의 상태로 복원하기 위한 추가 정보에 대하여 다음의 안전성 확보 조치를 해야 한다. 첫째, 개인정보의 안전성 확보 조치를 취하여야 한다. 둘째, 가명정보와 추가정보의 분리 보관. 다만, 추가정보가 불필요한 경우에는 추가정보를 파기해야 한다. 셋째, 가명정보와 추가정보에 대한 접근 권한을 분리하여야 한다. 다만, 「소상공인기본법」 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.[42]

개인정보처리자는 가명정보를 처리하고자 하는 경우에는 가명정보의 처리 목적, 제3자 제공 시 제공받는 자 등 가명정보의 처리 내용을 관리하기 위하여 다음의 사항 즉, 「가명정보 처리의 목적」, 「가명처리한 개인정보의 항목」, 「가명정보의 이용내역」, 「제3자 제공 시 제공받는 자」, 「그밖에 가명정보의 처리 내용을 관리하기 위하여 보호위원회가 필요하다고 인정하여 고시하는 사항」에 대한 관련 기록을 작성하여 보관하여야 한다.[43]

다른 한편 개인정보보호법 제28조의7과 신용보호법 제40조의 3이 정보주체에 대한 개인정보처리자의 여러 의무와 정보주체의 개인정보에 대한 권리를 가명정보에게는 광범위하게 배제하고 있다. 수집출처의 고지, 영업양도 등의 정보주체에의 고지, 개인정보 유출통지 등의 개별 정보주체에 대한 통지 내지

---

40) 김병필, 전게논문, 81면.
41) 개인정보보호법 제28의5, 신용보호법 제40조의 2 제8항
42) 개인정보보호법 시행령, 제29조의 5
43) 개인정보보호법 제28조의4 제2항, 개인정보보호법 시행령 제29조의5 제2항

고지의무가 모두 면제되고, 보유기간 경과 후 파기의무도 면제된다. 정보주체의 열람·정정·삭제·처리정지 등 개인정보 침해에 대한 손해배상청구권을 제외한 일체의 권리도 배제된다. 보유기간 경과 후 파기의무의 면제는 통계작성, 과학적 연구, 공익적 기록보존 등의 목적의 특성을 고려한 것이고, 그 이외의 의무와 권리의 배제는 가명처리를 한 결과 가명정보에 대하여 정보주체를 추적하기 어렵다는 점을 고려한 것이다.[44]

이 규정은 GDPR 제11조, 제12조 제2항의 영향을 받아 입법화 되었는데, 그 입법적 적절성에 대하여는 논란의 여지가 있다. 이에 관한 GDPR 규정은 컨트롤러가 정보를 보유하는 목적이 정보주체의 식별을 요하지 않는 경우 식별에 필요한 추가정보를 유지, 획득 또는 처리할 의무를 면한다. 컨트롤러가 정보주체를 식별할 수 있는 위치에 있지 아니함을 증명할 수 있을 때에는, 컨트롤러는 가능한 경우 정보주체에게 이를 고지하여야 한다. 이 경우 정보주체는 그의 식별을 가능하게 하는 추가정보를 제공하지 아니하는 한 제15조 내지 제20조의 권리를 행사할 수 없고, 각종 고지의무도 면한다고 규정하고 있다. 가명처리가 이러한 사유에 해당할 수 있으나 단순한 가명처리로 족하지 아니하고, 가명처리 결과 식별할 수 없게 되어야 하고, 그 경우에도 여전히 정보주체에게는 자신이 추가정보를 제공함으로써 특례를 배제할 가능성을 인정하고 있다. 이 점에서 우리 법은 GDPR에 비하여 정보주체의 권리를 지나치게 일률적으로 제한하고 있다는 비판도 가능할 것이다.[45]

6. 가명처리와 추가정보

추가정보는 개인정보의 전부 또는 일부를 대체하는 처리과정에서 생성 또는 사용된 정보로서 특정개인을 알아보기 위하여 사용·결합될 수 있는 정보를 말한다.[46] 대표적 예로 알고리즘, 매핑테이블 정보, 가명처리에 사용된 개인정보 등을 들 수 있다. 추가정보란 크게 2가지로 구분할 수 있다. 첫째, 원본 데이터 그 자체이다. 원본 데이터가 있다면 원래의 상태로 복원할 수 있음은 당연하다. 둘째, 전술한 가명처리의 비밀이다. 일방향 암호화를 이용하는 경우의 매핑 테이블 또는 쌍방향 암호화를 이용하는 경우의 암호키이다. 이 정보가 있다면 가명 값으로부터 원래 값을 복원해낼 수 있다. 따라서 가명정보를 활용하

44) 이동진, 전게논문, 168-169면.
45) 이동진, 전게논문, 169-170면.
46) 개인정보보호위원회, 가명정보처리 가이드라인, 7면.

는 연구자가 추가정보가 추가정보(원본 데이터나 가명정보 비밀)에 접근할 수 없도록 하여야 한다. 이처럼 연구자가 추가정보에 접근할 수 없는 상황에서 개인을 식별할 수 없도록 데이터를 변형하는 것이 가명처리이고 그 결과물이 가명정보이다.

가명정보는 추가정보가 있으면 개인정보를 알아볼 수 있는 정보이다. 그래서 가명정보도 여전히 개인정보에 해당한다. 개인정보보호법도 가명정보가 개인정보의 한 가지 유형이라는 점을 명시하고 있다. 하지만 가명정보는 가명처리를 통해 개인을 식별할 가능성을 크게 낮춘 것이다. 그래서 개인정보보호법은 가명정보 처리의 특례를 두고 있다. 통계작성, 과학적 연구, 공익적 기록보존 등의 목적을 위해서는 정보주체의 동의를 받지 않고 가명정보를 이용할 수 있고, 제3자에게 제공할 수도 있다.

7. 가명처리와 가이드라인

가명정보에 규제에 관한 법령상의 근거를 두지만, 가명처리의 핵심적 절차는 가이드라인의 형태로 규율된다. 가이드라인은 행정지도의 한 형태이다. 행정지도라 함은 행정기관이 그 소관 사무의 범위에서 일정한 행정목적을 실현하기 위하여 특정인에게 일정한 행위를 하거나 하지 아니하도록 지도, 권고, 조언 등을 하는 행정작용을 말한다.

행정지도에는 행정조직법적 근거가 필요하다는 데에 대해서 의견이 일치되어 있다. 문제는 행정작용법적 근거가 필요한가에 있다. 행정지도에는 행정작용법적 근거가 불필요하다는 것이 지배적 견해이다. 그 이유는 행정지도의 최대의 효용이 행정작용법적 근거가 결여되어 있음을 전제로 해서 행정기관이 새로운 행정수요에 기민하게 대응하여 행정책임을 수행하려는 데에 있으므로 행정지도에 일일이 행정작용법적 근거가 필요하다고 하면 행정지도의 기능은 발휘될 수 없기 때문이다. 그러나 행정지도 중에서도 규제적 행정지도에는 작용법적 근거가 필요하다는 견해가 유력하다.[47]

가명정보처리의 중요한 내용을 규율하는 개인정보보호위원회의 「가명정보처

47) 김철용, 「행정법 전면개정 제11판」(서울: 고시계사, 2021), 342면. 행정지도는 여러 기준에 따라 다양하게 분류할 수 있으나 대표적으로 그 기능에 따라 조성적 행정지도, 규제적 행정지도, 조정적 행정지도로 나눌 수 있다. 조성적 행정지도는 행정기관이 일정한 정책목적을 위하여 사인에 대하여 정보 등을 제공함으로써 사인의 활동을 조성해 주려는 행정지도이고, 규제적 행정지도는 사인의 활동을 규제할 목적으로 행하는 행정지도이다. 조정적 행정지도는 행정기관이 사인의 활동을 규제할 목적으로 행하는 행정지도이다.

리 가이드라인」은 그 내용상 규제적 행정지도, 조성적 행정지도의 일종으로 분류할 수 있다. 가명처리에 관하여 구체적 법률규정의 위임 없이 가이드라인 형식의 행정지도로 독자적으로 정하여 일반국민에게 부과한 것으로서 법률유보원칙과 조화되기 어려운 측면이 있다. 법률유보의 관점, 즉 행정권의 발동에는 작용법적 수권을 요한다는 관점에서 보면 가명처리에 관한 규율은 법률에 의한 행정이 아닌 행정지도에 의한 행정으로 변질되어 법치주의 실현이라는 측면에서 많은 문제점을 내재하고 있다. 이러한 법적 규율 체계는 법률유보의 범위에 관하여 중요사항유보설(본질성설)[48]을 취하고 있는 판례의 입장에서도 용인하기 힘든 부분이 있고, 헌법상 요구되는 위임입법에 있어서 수권법률의 명확성원칙의 관점에서도 개선이 요구되는 점이 적지 않다고 보여 진다.[49]

Ⅴ. 데이터 결합제도와 가명처리

개인정보처리자는 통계작성, 과학적 연구, 공익적 기록보존 등을 위하여 정보주체의 동의 없이 가명정보를 처리할 수 있다. 그러나 이러한 목적을 위한 것이라고 하더라도 서로 다른 개인정보처리자 간의 가명정보의 결합은 보호위원회 또는 관계 중앙행정기관의 장이 지정하는 전문기관이 수행하도록 하고 있다.[50]

1. 가명처리된 데이터의 결합

개인정보보호법 제20조 제1항은 「개인정보처리자가 정보주체 이외로부터 수집한 개인정보를 처리하는 때」에는 정보주체에게 개인정보 수집 출처 통지의 무를 이행할 것을 요구하고 있다. 데이터 결합과정에서 개인정보처리자가 결합된 가명정보를 제공받게 됨으로써 자신이 당초 가진 개인정보 이상의 정보를 추가로 가지게 되는 행위의 성질 역시 개인정보의 수집, 그중에서 정보주체 이외의 자로부터의 수집의 한 유형으로 파악할 수도 있다. 그러나 이렇게 해석하여 개인정보처리자가 개인정보보호법 제20조 제1항에 따른 수집 출처의 통지의무를 이행하는 행위는 개인정보보호법이 금지하는 재식별행위가 필연적으로

48) 대법원 2015. 8. 20., 선고, 2012두23808, 전원합의체 판결
49) 헌재 2005. 6. 30. 2005헌가1, 판례집 17-1, 796 [전원재판부]
50) 개인정보보호법 제28조의3 제1항

따를 수밖에 없을 것이다. 이러한 점을 고려하여 개인정보보호법은 가명정보에 대해 제20조의 적용을 배제하고 있다.[51]

가명정보의 주요한 활용 가능성 중의 하나는 다른 정보와 결합하여 활용하는 것이다. 예컨대 서로 다른 시점에 여러 기관에서 다양한 보건의료 데이터를 수집하여 보건의료 데이터를 결합함으로써 유용한 분석을 수행할 수 있다. 여러 의료기관의 데이터를 종합하면 환자가 경험하는 진료의 연속선상에서 의료서비스 전체 그림을 파악할 수 있게 된다. 보건의료 데이터와 다른 데이터를 결합하는 경우도 있을 수 있다. 이민자 정보와 보건의료 정보를 결합하면 이민 후의 건강상태를 분석할 수도 있다.

서로 다른 데이터셋들을 결합하려면 양자를 연결할 수 있는 식별자가 필요하다. 식별자는 단일한 필드일 수도 있고, 여러 필드를 합친 것일 수도 있다(이름+생년월일 등). 가명정보에 대한 데이터결합은 두 데이터셋에 대해 식별자를 암호화한 다음, 암호화된 식별자를 서로 대조하여 일치하는 항목을 찾는 것이다, 물론 두 데이터셋의 식별자에 대해 동일한 암호화 기법이 적용되어야 함은 당연하다. 그렇지 않다면 같은 식별자가 같은지 확인할 수 없을 것이다. 2020년 개정 개인정보보호법은 개인정보보호위원회나 관계 중앙행정기관의 장이 지정하는 기관에서만 가명정보의 결합을 수행하도록 규정하였다.[52] 가명정보의 결합을 정부가 지정한 기관에서만 수행할 수 있도록 한 것은 독특한 입법이다. 국내에서는 아직 이러한 기능을 수행할 민간 서비스가 충분히 활성화되지 못했다고 판단하였기 때문이라고 한다. 보건복지부는 2020년 10월 건강보험심사평가원, 국민건강보험공단, 한국보건산업진흥원 3곳을 보건의료 분야 결합전문기관으로 지정하였다.[53]

데이터 결합과 관련하여 유의할 사항은 데이터를 결합한 결과 개인식별 위험성이 증가한다는 점이다. 데이터셋 내에 간접 식별자 정보가 증가할수록 다양한 조합을 통해 개인을 식별할 가능성이 생겨나게 된다. 그래서 결합된 데이터는 결합전문기관 내에 설치된 별도의 공간에서 반출을 위한 추가적인 가명처리 또는 익명처리가 이루어져야 하고, 그 후 결합전문기관에 반출 신청을 하고 승인을 얻어야 한다.[54]

---

51) 제28조의7(적용범위) 가명정보는 제20조, 제21조, 제27조, 제34조제1항, 제35조부터 제37조까지, 제39조의3, 제39조의4, 제39조의6부터 제39조의8까지의 규정을 적용하지 아니한다.
52) 개인정보보호법 제28조의 3 제1항.
53) 김병필, 전게논문, 87면.
54) 개인정보보호법 제28조의3 제2항, 시행령 제29조의 3 제3항.

결합전문기관에 가명정보의 결합을 신청하려는 개인정보처리자는 보호위원회가 정하여 고시하는 결합신청서에 다음의 서류 즉 「사업자등록증, 법인등기부등본 등 결합신청자 관련 서류」, 「결합 대상 가명정보에 관한 서류」, 「결합목적을 증명할 수 있는 서류」, 「그밖에 가명정보의 결합 및 반출에 필요하다고 보호위원회가 정하여 고시하는 서류」를 첨부하여 결합전문기관에 제출해야 한다. 결합전문기관은 가명정보를 결합하는 경우에는 특정 개인을 알아볼 수 없도록 해야 한다.[55] 이 경우 보호위원회는 필요하면 한국인터넷진흥원 또는 보호위원회가 지정하여 고시하는 기관으로 하여금 특정 개인을 알아볼 수 없도록 하는 데에 필요한 업무를 지원하도록 할 수 있다. 결합전문기관은 결합 및 반출 등에 필요한 비용을 결합신청자에게 청구할 수 있다.

2. 가명처리된 데이터의 결합과정

서로 다른 개인정보처리자 간의 가명정보의 결합은 보호위원회 또는 중앙행정기관의 장이 지정하는 전문기관이 수행한다. 이 전문기관을 결합전문기관이라고 한다. 가명정보 결합ㆍ반출절차는 ① 결합신청, ② 결합 및 추가처리, ③ 반출 및 활용, ④ 안전한 관리 총4단계로 이루어진다.[56]

① 결합신청 : 결합신청자는 신청자간 결합신청에 필요한 사항의 협의, 결합신청서 작성 등 가명정보 결합에 필요한 사전 준비사항을 확인하고, 결합전문기관에 결합을 신청한다. 결합신청자는 결합전문기관과 결합일정, 전송방법 등을 협의한다.

② 결합 및 추가처리: 가명정보를 제공하는 결합신청자는 결합키 관리기간으로부터 결합키, 생성에 이용되는 정보값(Salt값)를 수신하여 결합키를 생성하고 결합신청 시 선택한 모의결합, 결합률 확인, 가명정보 추출 등이 완료되면 결합에 필요한 정보를 각 기관에 전송한다.

③ 반출 및 활용 : 결합정보 또는 분석결과 등을 반출하려는 경우, 결합전문기관에 반출을 신청한다.

④ 안전한 관리 : 결합정보를 이용하는 결합신청자는 반출한 결합정보를 당초 결합신청서 및 반출신청서에 기재한 목적에 따라 처리하고 안전조치 의무 등을 준수하여야 한다.

---

55) 개인정보보호법 제29조의3

56) 개인정보보호위원회, 가명정보처리 가이드라인, 2022. 4, 41면.

3. 결합 전문기관의 법적 성격

데이터 결합 제도가 도입됨으로써 결합된 데이터를 이용한 새로운 서비스 상품의 개발, 기존 서비스의 고도화를 추구하려는 많은 기업들에게는 새로운 도전과 기회를 찾아볼 수 있는 법적 기반이 조성되었다고 볼 수 있다. 이 경우 개인정보보호법은 데이터 결합을 반드시 중앙행정기관에서 지정된 전문기관에서 수행하도록 하고 있다. 이러한 데이터 결합과정에서 전문기관이 가지는 법적 지위에 대해 개인정보보호법은 특별한 언급을 하고 있지 않는다. 이와 관련하여 개인정보보호법은 개인정보 처리과정에서 행하여지는 개인정보 이전을 원칙적으로 개인정보의 제3자 제공이나 위수탁처리의 한 행위로 규정하고 있다, 그러나 동법은 제3자 제공에 해당하는 경우와 위수탁에 해당하는 경우의 구분에 관하여 아무런 언급이 없다. 개인정보의 제3자 제공에 관하여 개인정보보호법 제17조 제1항은 개인정보처리자는 다음 각 호의 어느 하나에 해당하는 경우에는 개인정보를 제3자에게 제공할 수 있다고 규정한다. 개인정보의 위수탁처리에 대하여 제26조 제2항은 「개인정보처리자가 제3자에게 개인정보의 처리 업무를 위탁하는 경우」, 「개인정보 처리 업무를 위탁받아 처리하는 자」라고만 규정하고 있을 뿐이다. 즉 어떤 경우가 개인정보의 제3자 제공인지 위수탁인지를 구별할 수 있는 기준을 전혀 제공하고 있지 않다. 이에 대하여 판례는 오래전부터 누구의 이익과 업무를 위한 이전인지에 따라 제3자 제공과 위수탁이 구별되고, 이러한 구분을 위해서는 다양한 요소를 고려하여 판단할 것을 요구하고 있다.[57] 대법원 판례의 입장을 기준으로 보면 개인정보보호법상 전문기관의 역할이 결합의뢰기관을 위해 가명정보를 결합제공하는 것이고, 결합된 데이터를 직접 이용하는 것은 상정하고 있지 않으므로 결합을 의뢰한 개인정보처리자에 대하여 수탁자로서 결합업무를 수행하는 것으로 보인다. 개인정보보호법이 데이터 결합, 그 중 외부적 결합은 반드시 전문기관을 통하여 하도록 법정하고 있다는 점에서 제26조의 특별규정이라고 볼 수 있다.

개인정보보호법 개정과정에서 전문기관으로 지정될 수 있는 기관에 민간기업이나 기관이 가능한지 여부에 대하여 논의가 있었지만 최종적으로 민간기업에 대해서도 요건만 충족하면 전문기관으로 지정될 수 있는 것으로 정리되었다. 다만, 개정법에 따른 데이터 결합 초창기에는 전문기관은 주로 비영리, 공공기관이 주류가 될 것으로 보인다.[58]

57) 대법원 2017. 4. 7., 선고, 2016도13263, 판결

4. 결합전문기관의 지정과 취소 및 감독

서로 다른 개인정보처리자 간의 가명정보의 결합은 보호위원회 또는 관계 중앙행정기관의 장이 지정하는 전문기관이 수행한다. 이러한 결합전문기관의 지정기준은 다음과 같다. 첫째, 보호위원회 고시에 따라 가명정보의 결합·반출 업무를 담당하는 조직을 구성하고, 개인정보 보호와 관련된 자격이나 경력을 갖춘 사람을 3명 이상 상시 고용할 것, 둘째, 보호위원회 고시에 따른 가명정보를 안전하게 결합하기 위하여 필요한 공간, 시설 및 장비를 구축하고 가명정보의 결합·반출 관련 정책 및 절차 등을 마련할 것, 셋째, 보호위원회 고시에 따른 재정 능력을 갖출 것, 넷째, 최근 3년 이내에 개인정보보호법 제66조에 따른 공표 내용에 포함된 적이 없을 것 등이다. 이러한 요건을 갖춘 때에는 상기의 보호위원회 또는 관계 중앙행정기관의 장은 지정신청서를 제출한 법인, 단체 또는 기관을 결합전문기관으로 지정할 수 있다.[59]

법령위반 시 결합전문기관은 취소될 수 있다. 보호위원회 또는 관계 중앙행정기관의 장의 결합전문기관 취소사유는 다음과 같다. 첫째, 거짓이나 부정한 방법으로 결합전문기관으로 지정을 받은 경우, 둘째, 결합전문기관 스스로 지정 취소를 요청하거나 폐업한 경우, 셋째, 제1항에 따른 결합전문기관의 지정기준을 충족하지 못하게 된 경우, 넷째, 결합 및 반출 등과 관련된 정보의 유출 등 개인정보 침해사고가 발생한 경우, 5. 그밖에 법 또는 이 영에 따른 의무를 위반한 경우 등이다. 이들 사유 중 첫째와 둘째 사유는 기속적 취소사유로 규정되어 있다.[60]

결합전문기관으로 지정된 이후에도 가명정보의 결합과 반출 및 적정하고 안전한 가명정보의 처리를 위하여 지속적 법적 통제·관리가 필요하다. 이를 위하여 보호위원회 또는 관계 중앙행정기관의 장은 결합전문기관을 지정한 경우에 해당 결합전문기관의 업무 수행능력 및 기술·시설 유지 여부 등을 관리·감독해야 한다.

---

58) 박광배, “데이터 결합제도의 효율적 활용 가능성”, 「데이터와 법」(서울: 박영사, 2021), 206면.
59) 개인정보보호법 시행령 제29조의 2 제1항
60) 개인정보보호법 시행령 제29조의 2 제5항

5. 전문기관을 통한 결합 · 반출심사

전문기관에서 결합된 데이터셋이 외부로 반출되기 위해서는 외부로 반출되기 위해서는 외부 전문가가 주축이 된 평가단에 의한 평가가 필수적이다. 구체적으로 개인정보보호법 시행령 제29조의3 제4항은 해당 전문기관에서 결합된 데이터를 반출하기 위해서는 반출심사위원회의 심사를 거쳐 해당 결합전문기관의 승인을 얻어야 가능한데, 이를 위해서 적정한 수준의 가명처리 내지 익명처리가 선행되어야 가능하다.[61] 개인정보보호법 시행령은 개인정보처리자 간 결합된 가명정보를 결합 전문기관으로부터 반출하기 위해서는 반출심사위원회의 승인을 얻어야 하고, 이때 「결합 목적과 반출 정보가 관련성이 있을 것」, 「특정 개인을 알아볼 가능성이 없을 것」, 「반출 정보에 대한 안전조치 계획이 있을 것」 등의 3가지 기준을 충족할 것을 요구하고 있다.[62]

가명정보처리 가이드라인[63] 에 따르면 반출심사는 결합된 정보를 반출할 자의 처리목적과 처리환경 등을 고려하여 이루어지므로 추가처리 시 가명정보처리수준은 결합된 정보를 활용하고자하는 자의 처리환경 등을 고려하여 판단하여야 하고, 익명정보로도 목적달성이 가능한 경우, 익명처리하여 반출하여야 한다. 결국 반출심사 역시 반출해서 이용할 자의 처리목적과 처리환경 등을 개별적으로 검토, 판단하여야 하고, 더 나아가 익명정보로 반출할 자의 목적달성이 가능한 지 여부도 판단하여야 하는 부담을 지게 되었다. 여기에 대한 판단을 위해 일률적으로 적용 가능한 기준을 정립하기는 쉽지 않을 것으로 보인다. 개인정보보호법 개정 전 과거 비식별조치 가이드라인 체제하에서는 전문기관은 외부 전문가의 추천, 기술적, 절차적 자문 등의 업무를 담당한 반면, 데이터 결합자체는 해당 결합신청기관에서 자체적으로 진행하게 함으로써 적정성 심사와 관련된 모든 업무가 데이터 결합신청기관과 적정성 평가단의 업무로 귀결이 될 수밖에 없었지만, 개정법 하에서는 전문기관이 결합 업무와 반출승인 관련 업무를 수행하는 만큼 과거 적정성 평가단의 업무를 일정 부분 떠맡아 평가단의 업무 부담을 줄일 여지를 마련한 점은 제도적 개선으로 보인다. 그러나 반출심사의 과정은 반출할 자의 처리목적과 처리환경을 검토해야 하기 때

61) 「신용정보의 이용 및 보호에 관한 법률」은 비록 이러한 반출심사위원회 심사절차를 명시하고 있지는 않지만, 동법 시행령 제14조의2 제3항 제5호에서 「데이터 전문기관은 결합된 정보집합물의 가명처리 또는 익명처리의 적정성을 평가한 후 적정하지 않다고 판단되는 경우 다시 가명처리 또는 익명처리하여 전달할 것」을 명시하고 있다.

62) 개인정보보호법 시행령 제29조의3 제4항

63) 개인정보보호위원회, 가명정보처리 가이드라인, 2022. 4, 56-57면

문에, 복잡하고 대규모의 개인정보처리가 불가피한 대기업의 데이터 결합이나 대량의 데이터 결합 등의 경우, 그 반출을 위한 적정성 심사가 상당한 부담이 될 수도 있다.[64]

6. 데이터 결합제도와 효율적 제고

GDPR의 경우 많은 부분에 있어서 데이터 결합제도에 대하여 위험기반 접근법(risk based approach)을 취하고 있는 반면, 우리 개인정보보호법은 규정기반 접근법(rule based approach)을 취하고 있다고 할 수 있다. GDPR은 특정 개인정보처리행위가 개인의 자유와 권리, 프라이버시에 실질적으로 어떠한 위험을 주는지를 고려하여, 위험의 정도에 상응한 적절한 규율과 보완책을 마련하도록 하는 위험기반 접근법을 취하고 있다고 할 수 있다.[65] EU 역외로의 개인정보의 이전(cross border transfer), 민감정보(special categories of personal data, criminal record 등)에 대한 특별한 예외적 조치 등 일부의 경우를 제외하면, 추상적인 요건이나 고려요소를 나열하고 그에 합당한 조치를 취하도록 하고 있다.

규정기반 접근법을 취하고 있는 우리 법제는 법 규정에서 추상적 표현이나 고려 요소의 나열을 최소화하고 개별규정에 대한 명확하고 엄격한 법령준수를 요구하고 있다.[66] 우리 법제는 기본적으로 개별 기업에 의한 위험평가의 요소를 최소화하고 드러난 위험과 그에 대응하는 조치를 유형화하여 법령준수를 요구하는 형태로 제도화하여 왔다. 이러한 접근은 위험평가(privacy risk assessment)의 개념보다는 개인정보처리자가 법규상의 의무사항의 준수를 요구하는 구조이다. 이러한 접근법을 택할 경우, 법령상의 의무만 준수하면 프라이버시 침해 위험이 해소될 것을 기대하는 구조로 법령은 진화할 수밖에 없고 기술진보와 환경변화에 따라 계속하서 새로운 의무 사항을 추가하는 구조로 발전할 수밖에 없을 것이다. 이는 결국 불필요하게 과도한 법령준수의무를 부담시키게 되고 개별 개인정보처리자가 데이터 처리환경에 적합한 비례성 있는 대응조치를 취할 수 있는 융통성을 없애는 결과를 초래하는 것으로 보인다. 이러한 과도한 규정기반 접근법으로는 복잡하고 급격히 진화하는 데이터 처리환경에 효율적으로 대처가 불가능하거나 데이터 처리 자체를 심각하게 제한하는

64) 박광배, 전게논문, 210-213면.
65) GDPR 제4조, 제35조
66) 개인정보보호법 제28조의3, 동법시행령 제29조의3

결과를 초래할 가능성도 있다. 이러한 점을 고려할 때, 외부 전문기관만을 통한 결합을 허용하는 제도를 고위험 데이터 결합에 한정하고 그 경우 심도 깊은 가명처리의 적정성 심사가 가능하도록 하는 제도개선도 검토할 필요가 있는 것 같다.[67]

Ⅵ. 맺음말

주관적 공권으로서 우리 헌법 해석상 인정되어 온 개인정보자기결정권은 정보주체가 타인이 보유하고 있는 자신의 정보에 접근하여 열람하거나 그 정보의 정정, 삭제, 차단 등을 요구함으로써 자신에 관한 정보에 통제할 수 있는 권리이다. 이를 보장하기 위해서 개인정보보호법에서는 정보주체의 동의를 개인정보처리의 원칙적인 요건으로 정하고 있는데 일반적인 공익보다 개인정보자기결정권을 우선시키는 것으로 평가된다.

동의요건 준수를 위하여 많은 사회적 비용이 발생한다. 이러한 점을 고려하여 개인정보보호법은 개인정보자기결정권을 보호하면서 동의요건 완화를 통한 개인정보의 이용확대를 위해서 가명정도제도를 적극적으로 법상 규정하게 되었다. 빅데이터와 인공지능이 산업과 사회발전의 핵심을 구성하는 새로운 미래를 준비하기 위해서는 대량의 데이터 수집과 분석이 필수 불가결한 것이고 가명정보제도가 가장 현실적 대안이라고 할 수 있다. 앞으로 개인정보보호를 강화하면서 가명정보를 통한 데이터활용 범위가 확대되도록 제도운용의 균형을 유지할 필요가 있다. 이러한 것을 실현하기 위하여 다음의 3가지 점은 추후 개선이 필요하다.

첫째, 추가정보는 개인정보의 전부 또는 일부를 대체하는 처리과정에서 생성 또는 사용된 정보로서 특정 개인을 알아보기 위하여 사용·결합될 수 있는 정보이다. 이러한 추가정보에 대한 법상 정의규정을 두고 있지 않다. 추가정보 범위를 적정하게 설정하지 않을 경우, 가명정보 의 적정한 범위 획정 상 혼란을 가져올 수 있다. 추가정보의 개념을 명확히 함으로써 제도 시행의 역기능을 사전에 예방할 필요가 있다.

둘째, 가명정보에 규제에 관한 법령상의 근거를 두지만, 가명처리의 핵심적 절차는 가이드라인의 형태로 규율된다. 행정권의 발동에는 작용법적 수권을 요

67) 박광배, 전게논문, 221면.

한다는 관점에서 보면 가명처리에 관한 규율은 법률에 의한 행정이 아닌 행정지도에 의한 행정으로 변질될 우려가 있다. 가명처리에 관한 법적 규율밀도를 높이는 것이 필요하다.

셋째, 데이터 결합제도에 대하여 과도하게 우리 개인정보보호법은 규정기반 접근법(rule based approach)을 취하고 있다. 부분적으로 위험기반 접근법(risk based approach)을 도입하여 복잡하고 급격히 진화하는 데이터 처리환경에 효율적으로 대처할 필요가 있다.

[참고문헌]

김창조, “정보주체의 개인정보자기결정권 보장과 개인정보의 활용”, 법학논고 75권 2021, 53-92면.
이동진, “가명정보의 특례”, 「데이터와 법」(서울: 박영사, 2021), 150-171면.
박광배, “데이터 결합제도의 효율적 활용 가능성”, 「데이터와 법」(서울: 박영사, 2021), 197-228면.
신수용, “보건의료 데이터의 비식별화”, 「보건의료와 개인정보」(서울: 박영사, 2021), 41-70면.
김병필, “가명처리와 가명정보의 이해”, 「보건의료와 개인정보」(서울: 박영사, 2021), 71-90면.
김강한, “가명화 개인건강정보 보호관련 기본권보장에 관한 연구”, 세계헌법연구, 제27권 2호, 2021.
박노형, 「개인정보보호법」(서울: 박영사, 2021)
신종철, 「개인정보보호법 해설」(서울: Jinhan M&B INC, 2020)
장상수, 「정보보호 및 개인정보보호 관리체계 개론」(서울: 생능출판, 2020)
손형섭, 「4차 산업혁명기의 IT · 미디어법 개정판」(서울: 박영사, 2020)
김종선 · 이혁기 · 정기정 · 정연돈, 「데이터의 익명화 개념이해와 최신기술동향」(서울: 휴먼사이언스, 2018)
이기혁 · 김원 · 홍준호, 「개인정보보호와 활용개론」(서울: 생능출판, 2017)
박노형 외 8인저, 「EU 개인정보보호법-GDPR을 중심으로-」(서울: 박영사, 2017)
개인정보보호위원회, 가명정보처리 가이드라인, 2022. 4.
최경진, 「개인정보판례백선」(서울: 박영사, 2022)
김철용, 「행정법 전면개정 제11판」(서울: 고시계사, 2021)
김동희 · 최계영, 「행정법 I 」(서울: 박영사, 2021)

Mark Stamp, Information security, principles and practice,third edition, Wiley 2022.
Irene Aldridge/Marco Avellaneda, Big data science, Wiley 2021.

[Abstract]

The Legal Control System over Pseudonymous Data

Kim, chang jo*

This thesis analyzes The legal control system for pseudonymous data. The research methods employed in this thesis include examining the development of legal systems, case laws and theories related to this problem in Korea, EU and Japan.

Personal Information Protection Act stipulates the consent of the data subject as a principle requirement for personal information processing. This legislative attitude is evaluated to give priority to the right to self-determination of personal information over the public interest. However, there are many social costs to comply with the consent requirements. In the Personal Information Protection Act, the definition of pseudonymous information was added recently in 2020, and by this law the personal information controller was able to process pseudonymous information without the consent of the data subject for statistical preparation, scientific research, and preservation of records in the public interest.

The legal system of the pseudonymous information is a system designed to protect the information subject's right to self-determination of personal information(Recht auf informationelle Selbstbestimmung) while enabling the use of personal information from the other side. In the process of operating the legal system of the pseudonymous information, it is necessary to establish practices of law enforcement so that these two values can be protected and developed at the same time.

Keywords : pseudonymous information, Personal Information, the consent of the data subject, the information subject's right to self-determination of personal information, Personal Information Protection Act

* Professor, Law School, Kyungpook National University