



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

생활과학 석사학위논문

빅데이터 환경에서의
개인정보 활용에 대한 소비자인식

2015년 8월

서울대학교 대학원

소비자학과

김 인 혜

빅데이터 환경에서의 개인정보 활용에 대한 소비자인식

지도교수 여 정 성

이 논문을 생활과학 석사학위논문으로 제출함
2015년 5월

서울대학교 대학원

소비자학과

김 인 혜

김인혜의 석사학위논문을 인준함

2015년 6월

위 원 장 _____ (인)

부위원장 _____ (인)

위 원 _____ (인)

국문초록

빅데이터 환경에서의 개인정보 활용에 대한 소비자인식

정보통신 기술이 발전하면서 기존과는 다른 형태의 개인정보, 즉 소비자가 온라인 사이트에서 활동하면서 적극적으로 생산하거나 기록되는 정보들이 증가하고 있다. 특히 사업자들에 의해 정보가 수집·활용되는 과정에서 개인의 프라이버시 침해와 정보 유출로 인한 피해가 발생할 가능성이 존재한다. 따라서 빅데이터 분석의 원자료인 정보를 생산하는 소비자들은 빅데이터 환경에서의 개인정보에 대해 어떻게 인식하고 있으며, 정보가 사업자에 의해 활용되는 것을 어떻게 받아들이고 있는지 연구할 필요가 있다.

본 연구에서 빅데이터 개인정보보호 가이드라인을 중심으로 빅데이터 환경에서의 개인정보를 크게 공개된 정보와 이용내역정보로 분류하고 개별 개인정보들에 대한 소비자의 인식을 살펴보았다. 즉 온라인 활동을 함으로써 생산된 개인에 관련된 정보들에 대한 소비자의 인식 수준은 어떠한지, 소비자집단에 따라 인식 수준에 차이가 존재하는지 탐구하였다. 또한 빅데이터 환경에서 생산된 개인정보를 사업자가 활용하는 것에 대해 소비자들이 어떻게 인식하고 있는지와 인식에 영향을 미치는 변수를 밝혀내었다.

조사전문업체인 엠브레인을 통해 온라인 사이트에 가입해본 경험이 있는 전국의 10대에서 50대 소비자 574명을 대상으로 조사한 자료를 분석하여 도출한 결과와 결론은 다음과 같다.

첫째, 전반적으로 소비자들이 빅데이터 환경에서의 모든 정보들을 자신의 개인정보라고 높게 인식하였는데, 본 결과는 기존 개인정보의 개념 및 범위를 확장할 필요가 있음을 시사한다. 세부적인 개인정보별 인식의 차이를 살펴보면, 자신이 공개한 이름과 성별 등의 일반적인 인적 사항

과 위치정보, 사이트에 접속했을 때 수집되는 실시간 기록을 개인정보로 인식하는 정도가 가장 높았다. 또한 쇼핑 내역과 같은 사이트 거래기록, 사진 및 동영상을 타 정보에 비해 자신의 개인정보라고 인식하는 점수가 높게 나타났다.

둘째, SNS나 온라인 쇼핑몰 이용률이 높고 일반적인 개인정보에 대한 관심과 우려가 높은 집단이 빅데이터 환경에서 생산된 개인정보에 더 민감하다는 사실을 밝혀내었다. 세부적인 결과를 살펴보면 남성에 비해 여성이, 50대에 비해 다른 연령대가 빅데이터 환경에서의 정보를 개인정보로 더 높게 인식하였다. 또한 개인정보를 보호하기 위한 행동 관련 점수가 높고, 개인정보 유출 경험이 있거나 유출이 되었는지 알 수 없다고 답한 소비자들이 정보를 자신의 개인정보라고 인식한 정도가 높았다.

셋째, 빅데이터 환경에서 이루어지는 사업자의 개인정보 활용에 대해 전반적으로 소비자가 동의하지 않았는데, 이는 소비자의 동의 없는 사업자의 개인정보 활용이 지양되어야 함을 의미한다. 특히 사업자가 소극적 절차를 통해 개인정보 활용에 대해 고지하였을 경우 동의하지 않는 정도가 높았다. 그러나 사업자가 적극적 절차를 통해 개인정보 활용에 대해 동의를 받았을 경우 소비자의 동의 정도가 높아졌다. 빅데이터 개인정보 보호 가이드라인에서 개인정보를 보호하기 위한 방안으로 강조한 비식별화 조치 즉, 사업자가 개인을 알아볼 수 있는 정보를 비식별화한다는 조건을 붙였음에도 동의 정도가 낮다는 것은 소비자가 사업자의 개인정보 활용을 기피하고 있음을 의미한다. 또한 개인정보를 활용하여 사업자가 소비자에게 맞춤형 광고와 같은 혜택을 제공할 때에도 동의 정도가 매우 낮았다.

넷째, 사업자의 개인정보 활용 여부에 대한 선택이 가능할 때 소비자들이 원하는 동의 방식을 살펴본 결과, 소비자들은 더 적극적인 선택권을 발휘하길 선호하였다. 온라인 사이트에 가입한 이후보다 가입하는 당시에 활용여부에 대한 선택권을 발휘하고자 하는 사람들이 더 많았다. 또한 사업자의 개인정보 활용 여부를 단 한 번 선택하고 그 선택이 이후 모든 정보 활용에 똑같이 적용되는 방식보다 정보가 활용될 때마다 선택

권을 가지길 선호하였다. 연구 결과는 소비자들이 개인정보에 대한 선택권을 더 적극적으로 보장받고자 함을 시사한다.

다섯째, 빅데이터 환경에서 사업자의 개인정보 활용에 대한 소비자의 동의를 높이기 위해서는 활용에 관련된 사실을 명확하게 전달하고, 소비자의 의사를 반영하는 가이드라인이 마련되어야 한다. 다중회귀분석 결과를 살펴보면, 개인정보 보호행동 점수가 높은 소비자들이 개인정보 활용에 동의하지 않는 것으로 나타났다. 또한 빅데이터 환경에서 생산되는 정보를 자신의 개인정보라고 인지할수록 활용에 동의하지 않았다. 하지만 개인정보가 활용되는 과정과 범위에 자신의 의사가 반영될 것이라고 생각하는 소비자가 그렇지 않은 소비자보다 기업이 정보를 활용하는 것에 대해 동의하였다.

연구 결과와 결론을 통해 다음과 같은 제언을 하고자 한다.

첫째, 소비자정책의 측면에서 소비자의 의견을 반영하여 사업자의 개인정보 활용에 대한 정보주체의 선택권을 강화하는 가이드라인이 수립되어야 한다. 소비자에게 사이트 가입 후 단순히 정보 활용에 대한 거부권을 행사할 수 있는 페이지를 제공하는 것보다 사이트 가입 단계에서 더욱 철저한 선택권을 부여해야 한다. 시대의 흐름에 맞추어 사업자의 빅데이터 수집 및 활용을 허용하고 이를 통해 가치를 생산할 수 있도록 독려하는 것이 중요하다. 하지만 정보주체인 소비자들이 개인정보 활용에 동의하는 수준이 매우 낮은 상황에서 사업자의 편의에 치우친 현 가이드라인은 수정되어야 한다.

또한 개인정보의 보호와 빅데이터 산업의 활성화 사이에 균형을 이루기 위해서는 지속적으로 빅데이터 환경에서 개인정보를 활용하는 사업자를 감시할 수 있는 감독기관을 설립할 필요가 있다. 개인정보를 활용하는 기업이 기술적·관리적 보호조치에 소홀하여 정보 유출이 발생하였을 시 기존과 다르게 강력한 대응이 이루어질 수 있는 제도를 도입할 필요가 있다.

둘째, 사업자는 변화하는 정보통신 환경에 발맞추어 빅데이터 산업에도 도전하는 현 시점에서 소비자와의 소통을 통해 동의를 이끌어내고 빅데

이터 환경에서 생산된 정보를 활용하여 가치를 창출하여야 한다. 빅데이터 환경에서 사업자는 개인정보를 수집하고 활용하는 방법과 범위에 대해 소비자에게 명확하게 전달하고, 그 내용을 이해시키려는 노력이 중요하다. 또한 개인정보를 보호하기 위해 필요한 안전한 기술적·관리적 보호조치를 도입하고 유지하는데 힘써 소비자의 신뢰를 형성하기 위해 지속적인 노력이 필요하다.

셋째, 소비자는 온라인상에서 자신이 생산하는 정보의 주체임을 자각하고, 정보에 대한 권리를 이해하고 실현하기 위해 노력해야 한다. 빅데이터 개인정보의 활용에 대해 정부와 기업의 조치만을 기다리는 수동적인 소비자가 아니라 적절한 조치를 요구할 수 있는 조직화된 힘을 보여주는 적극적 주체로 거듭나야 한다.

본 연구의 한계점과 후속연구를 위한 제언은 다음과 같다.

빅데이터 개인정보보호 가이드라인을 중심으로 개인정보의 종류를 구분하였으나 소비자가 개인정보를 생산하는 상황에 따라 빅데이터 환경에서 생산되는 개인정보의 분류가 달라질 수 있다. 따라서 상황과 맥락을 고려하여 생산되는 개인정보들을 분류하는 작업이 요구된다. 또한 빅데이터 개인정보가 지닐 수 있는 사회·경제·문화적 가치를 반영하여 이에 대한 소비자의 인식을 측정할 수 있는 보다 엄격한 척도와 변수 개발이 이루어져야 한다.

주요어 : 빅데이터 개인정보보호 가이드라인,
빅데이터 환경에서의 개인정보, 개인정보 인식,
개인정보 활용, 개인정보 활용 동의

학 번 : 2013-21527

목 차

제 1 장 서론	1
제 1 절 문제제기	1
제 2 절 연구목적	3
제 2 장 이론적 배경	4
제 1 절 빅데이터와 개인정보	4
1. 빅데이터의 개념과 특징	4
2. 빅데이터 환경에서의 개인정보	6
3. 빅데이터 관련 법과 제도	10
제 2 절 빅데이터 환경에서의 개인정보 활용에 관한 논의	21
1. 정보주체의 권리에 관한 논의	21
2. 정보처리자의 의무에 관한 논의	23
제 3 절 개인정보에 대한 소비자인식 관련 연구	25
1. 개인정보 침해에 대한 인식	25
2. 개인정보 및 개인정보 제공에 대한 인식	26
제 3 장 연구문제 및 방법	27
제 1 절 연구문제	27
제 2 절 연구방법	28
1. 조사대상 및 자료수집 방법	28
2. 변수의 정의 및 측정도구	29
3. 조사도구의 신뢰도	34
4. 분석방법	36
제 3 절 조사대상자의 일반적 특성	37

제 4 장 연구결과	39
제 1 절 빅데이터 환경에서의 개인정보에 대한 인식	39
1. 빅데이터 환경에서의 개인정보에 대한 인식수준	39
2. 빅데이터 환경에서의 개인정보에 대한 인식 차이	43
제 2 절 빅데이터 환경에서의 개인정보 활용에 대한 인식	48
1. 빅데이터 환경에서의 개인정보 활용에 대한 인식수준	48
2. 빅데이터 환경에서의 개인정보 활용에 대한 동의정도에 영향을 미치는 요인	55
제 5 장 결론 및 제언	59
제 1 절 요약 및 결론	59
제 2 절 제언	62
참고문헌	64
부 록	
첨부 1. 연구참여자용 설명서 및 동의서	72
첨부 2. 설문지	74
Abstract	81

표 목 차

[표 2-1] 개인정보의 유형	7
[표 2-2] Corby의 유비쿼터스 개인정보 분류	8
[표 2-3] 빅데이터 개인정보보호 가이드라인(안)의 주요내용	19
[표 3-1] 빅데이터 환경에서의 개인정보 유형과 종류	31
[표 3-2] 개인정보 활용에 대한 동의정도 측정 문항	33
[표 3-3] 조사도구의 신뢰도	35
[표 3-4] 조사대상자의 일반적 특성	38
[표 4-1] 빅데이터 환경에서의 정보를 개인정보로 인식하는 수준	40
[표 4-2] 정보 범주별 개인정보로 인식하는 정도의 차이 ·	42
[표 4-3] 인구통계학적 특성에 따른 인식 차이	44
[표 4-4] 개인정보 관련 행동과 경험에 따른 인식 차이 ...	47
[표 4-5] 개인정보 범주별 활용에 대한 동의정도	51
[표 4-6] 사업자의 개인정보 활용 범위에 따른 동의정도 ·	52
[표 4-7] 선호하는 개인정보 활용에 대한 동의 방식	54
[표 4-8] 개인정보 활용에 대한 동의정도의 영향요인	58

제 1 장 서론

제 1 절 문제제기

최근에는 SNS(소셜 네트워크 서비스)의 사용이 증가함에 따라 기존에 인터넷 사이트를 가입할 때 의무적으로 제공하였거나 거래 시 필수적으로 요구되었던 개인정보 외에도 소비자가 적극적으로 게시하는 사진, 동영상, 글 등의 정보가 대량으로 생산되고 있다. 디지털 환경에서 생성되는 데이터로 ‘규모가 방대하고, 생성 주기가 짧으며, 형태도 수치 데이터뿐만 아니라 문자와 영상 데이터를 포함하는 대규모 데이터’를 빅데이터라고 한다(정용찬, 2013). 과거에 비해 폭발적으로 양이 증가하였고, 종류가 다양해진 정보를 분석하는 기술을 활용하여 기업들은 사람들의 행동을 파악할 수 있을 뿐만 아니라 개인의 의견까지 취합하여 마케팅에 활용하고 있다. 소비자들이 스스로 인식하지 못하는 사이에 자신이 작성한 글이나 사이트를 이용한 내역 등이 사업자에 의해 수집·분석되고 있는 것이다.

빅데이터 산업이 증대됨에 따라 빅데이터 분석에 활용되는 개인정보에 대한 관심도 커지고 있다. 빅데이터 이전에는 유비쿼터스 개인정보라는 용어로 정보가 분류되기도 하였는데(Corby, 2002), 온라인 아이덴티티 정보 및 동태정보와 함께 새롭게 생성되고 있는 SNS 정보까지 포함하여 빅데이터 개인정보라고 일컬어진다(성준호, 2013). 빅데이터 환경에서 생산된 개인정보를 활용하여 경제적 이익을 취득하는 사업자를 기준으로 빅데이터 산업에 대한 분석이 주로 이루어지고 있다. 또한 정작 정보의 생산 주체인 소비자들이 이 정보를 어떻게 생각하고 있는지, 즉 기존과는 다른 정보를 어느 정도 자신의 개인정보로 인식하는지에 대한 소비자 인식 연구는 부족한 실정이다.

빅데이터 환경에서 기업의 개인정보 활용에 대한 사회적 합의를 도출

하기 위해 방송통신위원회는 2013년 12월 ‘빅데이터 개인정보보호 가이드라인’을 발표하였다. 이후 지난해 12월에 제시된 가이드라인 수정안은 정보통신서비스제공자가 빅데이터 분석에 개인정보를 활용하는데 있어서 어떤 조치를 취해야하는지 명시하고 있다. 가이드라인에서는 ‘공개된 정보’와 ‘이용내역정보’에 대한 정의를 통해 개인정보의 포괄적인 범주만 정하였다. 또한 초안에서 논란이 되었던 부분, 즉 사업자가 개인정보를 사용할 때 소비자의 동의를 얻지 않아도 된다는 부분은 수정되어 비식별화 조치가 추가되었다. 하지만 여전히 비식별화 조치를 취한다면 소비자의 사전동의를 구하지 않고 추후에 정보활용 거부권을 제공한다는 점에서, 가이드라인은 정보를 생산하는 주체인 소비자의 입장보다는 정보를 사용하는 사업자의 입장을 더 배려하고 있다.

지금까지는 원자료인 빅데이터를 생산하는 소비자보다 주로 이를 수집·활용하는 기업의 입장에서 빅데이터의 가치가 논의되어졌다. 그렇다면 소비자 입장에서 중요하게 다루어야 할 빅데이터 환경에서의 개인정보는 무엇을 의미하는지 살펴보고자 한다. 즉 온라인 활동을 함으로써 생산된 개인에 관련된 정보에 대한 소비자의 입장을 밝혀내고자 한다.

구체적으로 본 연구는 소비자들이 빅데이터 환경에서 생산된 정보를 자신의 개인정보라고 인식하는 수준에 대해서 알아보하고자 한다. 또한 사업자가 온라인에서 수집한 정보를 활용하는 것에 어느 정도 동의하는지 그 수준을 살펴보고, 소비자는 어떤 방식을 통해 개인정보의 활용에 있어서 자신의 선택권을 발휘하고자 하는지 밝혀내고자 한다.

제 2 절 연구목적

본 연구는 소비자가 빅데이터 환경에서 생산된 정보를 자신의 개인정보라고 생각하는 정도에 대해 살펴보고, 소비자 집단별로 인식의 차이가 발생하는지 알아보고자 한다. 또한 사업자가 빅데이터 환경에서 개인정보를 활용하는데 소비자가 동의하는지 파악하고, 이에 영향을 미치는 변수를 밝히고자 한다. 본 연구의 목적은 다음과 같다.

첫째, 빅데이터 환경에서 생산되는 각각의 정보에 대해 소비자가 자신의 개인정보라고 인식하는 수준을 알아보고자 한다. 기존 연구에서는 다루어지지 않았던 공개된 정보와 이용내역정보 등 각 범주에 속한 개별 개인정보에 대한 소비자의 인식을 살펴보고자 한다. 본 연구는 정보주체인 소비자가 인식하는 개인정보의 범위가 빅데이터 분석에 사용되는 정보까지 포함하여 확대되었음을 밝혀, 보다 확장된 범위의 개인정보 논의가 필요함을 주장하고자 한다.

둘째, 빅데이터 환경에서의 개인정보를 사업자가 활용하는데 있어 소비자가 동의하는 정도를 알아보고자 한다. 개인정보 범주별뿐만 아니라 사업자가 취하는 조치에 따라서 소비자가 개인정보 활용에 동의하는 정도에 차이가 발생하는지 살펴보고자 한다. 또한 이에 대한 영향요인을 밝혀 사업자가 정보주체의 진정한 동의를 바탕으로 빅데이터 산업을 활성화할 수 있는 방안을 도출하는데 기여하고자 한다.

셋째, 소비자가 사업자의 개인정보 활용에 대해 선택을 할 수 있다면, 어떤 방식을 선호하는지 살펴봄으로써 이후 빅데이터 개인정보보호 가이드라인이 수정되는 과정에서 실질적으로 소비자의 의사가 반영될 수 있는 방안을 제시하는데 도움이 되고자 한다. 빅데이터 환경에서 생산된 개인정보에 대한 소비자의 인식 연구를 통해 가이드라인을 수정하는 현 시점에서 소비자의 입장을 대변할 수 있을 것으로 기대한다.

제 2 장 이론적 배경

제 1 절 빅데이터와 개인정보

1. 빅데이터의 개념과 특징

인터넷의 보급과 함께 스마트폰 등 스마트기기의 보급이 확대되면서 기존에는 상상할 수 없었던 막대한 양의 데이터가 수집되고 분석에 활용되고 있다. 이처럼 ‘전통적인 데이터베이스 소프트웨어 툴을 이용해 수집, 저장, 관리 및 분석할 수 있는 규모를 넘어서는 데이터’가 바로 빅데이터이다(McKinsey Global Institute, 2011).

빅데이터의 특징은 기존의 3V 즉, 규모(Volume), 속도(Velocity), 종류(Variety)에 최근 가치(Value)가 추가되어 4V로 설명될 수 있다. 먼저, 빅데이터는 규모(Volume)가 매우 방대하여 수집·저장·관리·분석할 수 있는 데이터가 수십 테라바이트를 넘어서 페타바이트에 이른다. 단순히 데이터의 물리적인 양이 증가한 것뿐만 아니라 이를 분석·처리하는 과정에서 발생하는 데이터도 증가하였다(성준호, 2013).

방대한 양의 정보가 수집될 수 있다는 것뿐만 아니라 축적된 정보가 신속하게 처리될 수 있는 것이 빅데이터의 두 번째 특징으로, 바로 빠른 속도(Velocity)이다. 수많은 정보가 실시간으로 수집되고 처리될 수 있는 기술적 특징을 지니고 있는 것이다(성준호, 2013).

수집하고 처리할 수 있는 데이터의 종류(Variety)가 다양해진 것이 빅데이터의 세 번째 특징이다. 기존에 수집하는 일반적인 정형화된 데이터(Structured data)뿐만 아니라 사진, 오디오, 비디오 형식의 데이터나 로그 데이터와 같은 비정형화된 데이터(Unstructured data), 고정된 시스템에 저장되어 있지 않은 XML 혹은 HTML-tagged text 등을 포함하는 반정형화된 데이터(Semi-structured data)가 빅데이터에 포함된다(ITU, 2013).

빅데이터가 주목받고 있는 가장 중요한 이유는 다양한 데이터가 처리

되어 가치(Value)를 생성한다는 특징 때문이다. 빅데이터 안에는 서로 다른 경제적 가치가 내재되어 있는데, 이 중 가치 있는 정보를 파악하는 것이 중요해지고 있다(Oracle, 2012). 빅데이터 시장 규모를 살펴보면 2015년에 약 384억 달러에 이를 것으로 예상되며, 안정적인 성장세를 통해 2017년에는 최대 501억 달러로 성장할 것이라고 한다(Wikibon, 2014).

빅데이터의 가치를 중심으로 사업자의 관심이 모이고 있다. 사업자는 소비자에 관한 정보를 최대한 수집, 분석하고자 하며 이때 소비자의 프라이버시 문제가 제기된다(김혜선 등, 2004). 온라인을 통해 대량으로 빠르게 생산되는 소비자들의 구매기록, 방문한 사이트기록 등 로그데이터 정보는 소비자의 사적인 성향과 소비습관 및 행동을 쉽게 파악할 수 있는 단서가 된다. 빅데이터 처리 기술의 발달로 인해 기존에는 불가능했던 방대한 양의 소비자 정보가 수집·분석되는 과정에서 소비자 개개인을 식별할 수 있는 프라이버시 침해가 발생할 수 있는 것이다.

또한 빅데이터 분석을 위해 수집된 정보들의 통제가 과연 정보주체에 의해 이루어질 수 있는지에 대한 문제가 존재한다. 정보주체의 사생활이 침해될 수 있다는 논의에서 확대되어 자신에 관한 어떤 정보를 사업자가 어떤 방식으로 수집·분석하는지 정보주체가 스스로 결정할 수 있는가, 즉 정보에 대한 자기결정권 문제로 발전한 것이다(Stone et al., 1983; 김종길, 2009). 그렇다면 소비자 입장에서 중요하게 다루어야 할 빅데이터 환경에서의 개인정보는 무엇인지 구체적으로 살펴보고자 한다.

2. 빅데이터 환경에서의 개인정보

일반적인 개인정보에 대한 정의를 살펴보면 다음과 같다. 개인정보 관련법에서 명시하고 있는 정의를 보면, 「개인정보보호법」에서는 ‘살아 있는 개인에 관한 정보로서 성명, 주민등록번호 및 영상 등을 통하여 개인을 알아볼 수 있는 정보(해당 정보만으로는 특정 개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 알아볼 수 있는 것을 포함한다)’를 개인정보라고 한다(제2조). 「정보통신망 이용촉진 및 정보보호 등에 관한 법률(이하 정보통신망법)」에서는 ‘생존하는 개인에 관한 정보로서 성명·주민등록번호 등에 의하여 당해 개인을 알아볼 수 있는 부호·문자·음성·음향 및 영상 등의 정보(당해 정보만으로는 특정 개인을 알아볼 수 없는 경우에도 다른 정보와 용이하게 결합하여 알아볼 수 있는 것을 포함한다)’라고 정의한다(제2조 제6호).

법에서 정의하고 있는 개념 외에도 개인정보를 개념화하려는 연구가 이루어지고 있다. 온라인 상 수집되는 개인정보 중 사업자가 수집하는 정보를 기준으로, 홈페이지에서 회원가입 시 필수적으로 수집하는 항목을 개인정보로 정의하기도 한다(김민천, 2008). 개인정보를 보다 넓은 의미로 해석하기도 하는데, 디지털 시대의 개인정보는 네트워크상에 존재하는 개인에 관한 정보와 개인의 활동 및 사생활에 관한 자료들, 개인이 게시한 자료까지 포함한다(조성은·유지연, 2012).

개인정보에 대한 정의만큼이나 개인정보를 유형화하려는 시도도 다양하다. 방송통신위원회와 한국인터넷진흥원에서 발간한 가이드에서 분류한 개인정보의 유형과 구체적인 예시는 [표 2-1]과 같다. 크게 6가지로 분류되는데 먼저, 성명이나 주민등록번호, 주소와 같은 일반적인 인적사항과 지문, 홍채, 건강·의료 등에 관한 신체적 정보, 정치적 성향, 기호 등의 정신적 정보, 카드·계좌 번호 등의 재산적 정보, 학력, 직장 등의 사회적 정보, 그 외 사업자의 인터넷 사이트 가입이나 서비스 이용 시 생성되는 로그 기록, 구매 내역, 쿠키, 통화 내역, 위치 정보 등의 기타 정보가 포함된다(방송통신위원회·한국인터넷진흥원, 2011).

[표 2-1] 개인정보의 유형

유형	개인정보의 예	
인적 사항	성명, 주민등록번호, 주소, 본적지, 전화번호 등 연락처, 생년월일, 출생지, 이메일 주소, 가족관계 및 가족구성원 정보 등	
신체적 정보	신체정보	얼굴, 지문, 홍채, 음성, 유전자정보, 키, 몸무게 등
	의료·건강정보	건강상태, 진료기록, 신체장애, 장애등급, 병력 등
정신적 정보	기호·성향정보	도서·비디오 등 대여기록, 잡지구독정보, 물품구매내역, 웹사이트 검색내역 등
	내면의 비밀 등	사상, 신조, 종교, 가치관, 정당·노조 가입여부 및 활동내역 등
재산적 정보	개인금융정보	소득, 신용카드번호, 통장계좌번호, 동산·부동산 보유내역, 저축내역 등
	신용정보	신용평가정보, 대출 또는 담보설정 내역, 신용카드 사용내역 등
사회적 정보	교육정보	학력, 성적, 출석상황, 자격증 보유내역, 상벌기록, 생활기록부 등
	법적정보	전과·범죄 기록, 재판 기록, 과태료 납부내역 등
	근로정보	직장, 고용주, 근무처, 근로경력, 상벌기록, 직무평가기록 등
	병역정보	병역여부, 군번, 계급, 근무부대 등
기타	전화통화내역, 웹사이트 접속내역, 이메일 또는 전화메시지, 기타 GPS 등에 의한 위치정보 등	

참고: 방송통신위원회·한국인터넷진흥원, 2011.

개인정보를 이용자의 입장에서 즉, 사업자 중심에서 벗어나 정보 주체의 시각에서 분류한 연구도 등장하였다. 방송통신위원회 및 한국인터넷진흥원에서 제시한 개인정보를 바탕으로 온라인 기록정보, 통용인적정보, 라이프 스타일 정보 그리고 전용 신원정보 등 네 가지로 분류하였다(김민성·김성태, 2014).

Corby(2002)는 유비쿼터스 환경에서의 개인정보를 개인의 실재를 파악할 수 있는 시간과 공간의 개념을 포함하여 정태정보, 동태정보, 파생정보로 분류하였다(나중연, 2009; 조동기·김성우, 2003). 정보 주체의 행동의 흐름이 담긴 통시적 데이터와 실시간 데이터 등 동태정보와 기존의 정보들이 통합되어 발생하는 분석데이터 등이 파생정보이다.

[표 2-2] Corby의 유비쿼터스 개인정보 분류

데이터 유형		하위 유형	
정태 정보	아이덴티티	오프라인	생체정보, 경제정보, 법적정보, 사회적정보, 관계적 정보, 부동산관련 정보
		온라인	디지털 아이덴티티
	자산	유형자산	재산
동태 정보	통시적 데이터	낮은 수준	거래내역, 여행기록, 통화내역 등
		높은 수준	유비쿼터스 환경에서 수집된 로그 데이터(시간, 장소)
	실시간 데이터	유비쿼터스 환경에서 수집된 데이터	
파생 정보	분석데이터	시간 흐름에 따라 수집되고 분석된 데이터	
	통합데이터	다른 데이터와 결합된 개인정보	

기업들이 마케팅을 위해 온라인상에서 수집하는 사이트 접속기록이나 구매내역, 접속시간대 등과 같은 행태정보와 IP, 스마트폰 위치 등과 같은 위치정보 등을 소비자들의 일반적 정보와 결합하여 분석한 소비패턴 등이 파생정보인 것이다(정상조 외, 2010). 이러한 맥락에서 개인의 ID를 중심으로 개인의 성향 등을 포함한 파생정보와 개인의 행동 추이 등의 동태정보, 개인이 이용한 서비스 정보 또한 디지털 시대에서 중요한 개인정보로 포함된다(Cartrysse, 2003).

유비쿼터스 개인정보, 디지털 정보 등 다양한 명칭으로 분류된 개인정보의 공통점은 모두 인터넷의 발달로 인해 중요성이 부각된 개인의 행동 및 생각 정보를 포함한다는 것이다. 정보통신기술의 발달로 정태 정보의 축적이 용이해지고 방대한 동태 정보의 생산이 가능할 것이라는 예측은 근 10년 전부터 끊임없이 예견되었다(조동기·김성우, 2003). 실제로 빅데이터 활용 기술이 발달하며 예측은 현실화되었고, 기존의 보호영역에 속한 온라인 사이트 가입 시 제공했던 개인정보와는 다른, 사이트 가입 후 정보 주체가 제공하는 글, 사진 및 동영상 등이 광범위하게 수집·분석되고 있다.

개인정보에 포함되기 위해서는 개인을 식별할 수 있는 정보, 즉 해당 정보와 다른 정보를 결합하여 특정인을 다른 사람과 구분·구별할 수 있는 정보여야 한다. 기존에는 포함되지 않았던 방대한 양의 정보가 수집되고 결합되기 때문에 단순히 이름이나 주민등록번호를 제거하였다고 해서 개인에 대한 식별성이 없어졌다고 할 수 없다. 개인정보를 익명화하고 나머지 정보로 다른 사람과 구별될 수 있다는 문제의 여지가 남아있다(이창범, 2013). 따라서 어디까지를 개인정보에 포함할 수 있는지 보다 더 나아가 개인정보에 포함할 수 없는 정보를 명확하게 구분할 수 없는 시대가 도래한 것이다.

이미 구글, 아마존 등 기업들은 소비자에 대한 수많은 데이터를 분석하여 맞춤형 광고와 같은 서비스를 통한 마케팅 활동에 주력하고 있다. 사업자가 소비자의 데이터를 수집하는 과정에서 소비자의 동의 없이 무단으로 이루어지고 데이터가 활용될 수 있는 상황이 발생할 가능성이 증

대되고 있다. 따라서 소비자가 개인정보라고 인식하는 정보의 범위가 기존에서 어느 정도 확장되었는지 살펴보고, 이에 대한 보호를 논하는 것이 시급하다.

빅데이터 시대의 개인정보는 이전에 개인적인 영역으로 여겨졌던 사적인 생각과 행동들도 포함하며, 정보가 온라인상에 기록된 후에는 개인의 통제를 벗어나게 되므로 보다 광범위하게 정의되어야 하는 개념이다. 이런 맥락에서 본 연구에서는 빅데이터 환경에서의 개인정보에 소비자가 온라인 활동을 함으로써 생산된 개인과 관련된 정보를 포함하여 살펴보고자 한다.

3. 빅데이터 관련 법과 제도

1) 국내 개인정보관련법 고찰

국내 개인정보보호 관련법은 영리 또는 비영리를 목적으로 개인정보를 처리하는 모든 공공·민간·단체·개인 등을 대상으로 개인정보 보호에 관하여 규정하는 일반법인 「개인정보보호법」과 공공부문 및 민간부문의 수많은 개별법으로 이루어져있다. 즉 개인정보보호 법제는 정보통신, 금융·신용, 의료, 공공행정, 교육 등 개별 분야에 따라 각기 다른 법률이 제정되고 시행된다. 이처럼 민간 부문은 분야별로 다양한 개인정보보호 관련 법률이 존재하는데, 민간 영역에서 규제범위가 가장 넓고 개인정보의 취급단계별로 상세한 보호원칙을 두고 있는 것이 「정보통신망법」이다(행정안전부·한국인터넷진흥원, 2011). 따라서 본 연구에서는 「개인정보보호법」과 「정보통신망법」에서 다루고 있는 사업자의 개인정보 처리방침과 개인정보 취급방침을 살펴보고자 한다.

먼저, 「개인정보보호법」에서 다루고 있는 개인정보처리자란 “업무를

목적으로 개인정보파일을 운용하기 위하여 스스로 또는 다른 사람을 통하여 개인정보를 처리하는 공공기관, 법인, 단체 및 개인 등”을 말한다(제2조 제5호). 개인정보처리자의 다양한 의무를 규정하고 있으며, 개인정보의 안전한 관리를 위해 규정하고 있는 조치로서 개인정보 처리방침이 명시되어 있다(제30조).

개인정보 처리방침에 명시하여야 할 사항은 개인정보의 처리 목적과 처리 및 보유기간이 있으며, 해당되는 경우 제3자 제공 또는 위탁에 관한 사항을 공개하여야 한다. 또한 정보주체의 권리·의무 및 그 행사방법에 관한 사항을 기재하고, 처리하는 개인정보의 항목, 개인정보의 파기에 관한 사항, 개인정보의 안전성 확보조치에 관한 사항 등을 명시해야 한다.

다음으로, 「정보통신망법」에서는 “전기통신사업자와 영리를 목적으로 전기통신사업자의 전기통신역무를 이용하여 정보를 제공하거나 정보의 제공을 매개하는 자”인 정보통신서비스 제공자를 규제한다(제2조 제1항 제3호). 개인정보보호법 상 개인정보 처리방침과 같은 개념이지만 다른 명칭으로 규정하고 있다. 「정보통신망법」에는 소비자의 알 권리와 개인정보 자기결정권을 보장하기 위해 개인정보 취급방침에 대한 공개를 명시하고 있다(제27조의 2).

「정보통신망법」 상 개인정보 취급방침의 규정은 「개인정보보호법」에서 명시하고 있는 규정보다 더 상세하며 구체적이다. 사업자는 소비자의 개인정보를 취급하는 경우 개인정보 취급방침을 정하여 소비자가 언제든지 쉽게 확인할 수 있도록 공개하도록 명시하고 있으며 총 7가지 항목을 포함하도록 한다. 먼저 개인정보 수집·이용 목적과 수집하는 개인정보의 항목 및 수집방법을 공개하고, 목적에 필요한 만큼만 최소한으로 수집하여야 한다. 사업자가 소비자로부터 획득한 개인정보를 제3자에게 제공하는 경우 제공받는 자의 성명과 이용목적 및 제공하는 항목을 공개해야 한다. 개인정보 보유 및 이용 기간과 파기절차 및 파기방법, 보존하여야 하는 경우 그 근거와 항목에 대하여 명시해야 하며, 원칙적으로 개인정보 수집 및 이용목적이 달성된 후에는 관련 정보를 지체 없이 삭제

하여야 한다. 개인정보 취급위탁을 하는 업무 내용과 수탁자를 공개해야 한다. 또한 이용자 및 법정대리인의 권리와 그 행사방법에 대하여 명시해야 하는데, 정보주체가 자신의 개인정보를 열람, 정정, 삭제, 동의·철회할 수 있는 권리의 존재하는 것과 행사할 수 있는 방법에 대한 것이다. 다음으로 인터넷 접속정보파일 즉, 쿠키, 로그기록 등의 개인정보를 자동으로 수집하는 장치의 설치·운영 및 그 거부에 관한 사항을 공개해야 한다. 마지막으로 개인정보관리책임자의 성명 또는 개인정보보호 업무부서의 명칭 및 연락처를 기재하여야 한다.

2) 국외 개인정보관련 제도 고찰

(1) 미국

미국은 연방정부기관에서 개인정보를 보호하기 위해 보유하고 있는 ‘프라이버시법(Federal Privacy Act, 1974)’을 근거로 공공부문에 대한 규제를 하고 있으며, 민간부문에서는 시장의 자율규제를 원칙으로 하고 있다. 이후 1978년 금융프라이버시권리법, 1986년 전기통신보호법, 1988년 컴퓨터자료의 상호비교 및 프라이버시보호법, 1994년 전기통신프라이버시법, 1996년 통신법 및 의료기록비밀보호법 등을 제정하여 시대의 흐름과 신기술로 인한 개인정보의 위협에 대응하고 있다(권현영, 2012).

최근 2010년 ‘인터넷 프라이버시 법안(Internet Privacy Act)’과 2011년 ‘상거래 프라이버시 권리헌장(Commercial Privacy Bill of Right Act)’ 등을 통해 포괄적으로 인터넷 상에서 발생하는 개인정보를 보호하기 위한 법률을 제정했지만 근본적으로 정부의 간섭을 최소화하고 기업의 자율규제를 중시하는 기초를 유지하고 있다(김규남, 2014). 또한 의료정보, 통신정보, 아동정보 등을 제외한 개인정보를 처리하는데 옵트아웃(Opt-out) 방식을 허용하고 있어 빅데이터 활용에 유리한 법적 환경을 갖추고 있다(이창범, 2013).

하지만 최근 미국에서도 디지털 경제의 발전과 함께 소비자의 프라이

서비스가 위협받고 있는 상황을 개선하고 개인정보를 보호하기 위해 변화하고 있다. 정보 주체의 권리를 보장하고 온라인 서비스 업체에게는 책임을 강화하고 있는 예로 2012년 오바마 정부에서 발표한 ‘온라인 프라이버시 프레임워크(Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy)’와 FTC에서 발표한 프라이버시 보호 가이드인인 ‘소비자 프라이버시 보호(Protecting Consumer Privacy in an Era of Rapid Change)’가 대표적이다(나중연, 2013).

오바마 정부의 프레임워크의 핵심은 ‘소비자 프라이버시 권리장전(Consumer Privacy Bill of Rights)’을 통해 소비자의 개인정보보호와 기업에게 올바른 투자를 위한 7가지 원칙을 제시한 것이다. 권리장전에는 1) 개인정보의 자기결정권, 즉 개인정보의 수집 및 이용 과정에서 정보 주체의 선택권과 동의·철회권, 2) 소비자들이 프라이버시 및 보안성에 쉽게 접근할 수 있는 투명성, 3) 정보 주체가 자신의 정보를 제공할 당시의 조건 그대로만 정보가 사용되는 목적 외 이용·제공 금지, 4) 소비자들의 개인정보를 보호하기 위한 기술적 조치를 유지하고 관리하는 정보보안, 5) 정보 주체가 개인정보를 열람 및 정정할 수 있는 요구권의 보장과 기업의 정확성, 6) 기업들이 개인정보를 최소한으로 수집·보유하도록 제한, 7) 기업이 권리장전을 준수하는 책임에 관한 내용이 담겨있다(김선남·이환수, 2014).

FTC에서 발표한 ‘소비자 프라이버시 보호’는 소비자들이 자신의 정보 수집과 활용에 대한 통제권을 가질 수 있도록 기업이 취해야할 권고사항을 제시하고 있다. 기업의 프라이버시 보호활동, 소비자의 선택에 있어서 단순화, 프라이버시의 공시 및 소비자의 정보접근에 대한 투명성과 같은 3가지 원칙이 담겨있다(나중연, 2013). 기업의 자율규제를 유지하면서도 계속해서 노력을 요구함으로써 데이터의 수집 및 활용에 대한 소비자의 신뢰를 바탕으로 빅데이터 산업의 발전을 도모하고 있는 것이다.

(2) 유럽연합

유럽은 인권보호 측면에서 개인정보보호 관련 법제를 정비하여 개인의 프라이버시를 인간의 기본권으로 보장하고 있다. 1995년 유럽연합이사회는 데이터 보호 지침(Data Protection Directive)을 제정하였는데, 이는 유럽을 포함하여 전 세계의 개인정보보호 관련법에 영향을 주었다.

데이터 보호 지침에서는 정보주체에게 고지하여야 할 사항을 규정하고 있는데, 이는 개인정보 취급방침과 유사하다. 회원국은 관리자나 그 대리인이 개인정보를 수집할 경우, 해당 개인정보의 주체에게 관리자 및 그 대리인의 신원, 개인정보 처리의 목적, 정보의 수령인, 필수 또는 선택적 동의인지 고지해야 한다. 만약 동의하지 않을 경우에는 이로 인해 가능한 결과와 정보주체와 관련된 정보의 접근권 및 정정권이 있음을 고지해야 한다. 또한 정보가 수집되는 특정 상황과 공정한 정보처리의 보장과 관련하여 필요한 정보를 알리고, 정보주체로부터 정보를 획득하지 않은 경우 관리자 또는 그 대리인이 개인정보를 처리하는 시점과 제3자에게 공개된다면 최초로 정보가 공개되기 전에 명시된 고지사항을 정보주체에게 알려야 한다(전은정 외, 2012a).

이후 세계화와 정보통신의 발달로 국제적 협력을 통한 온라인 개인정보보호에 대한 필요성이 제기되어 유럽연합은 1995년 제정한 데이터보호 지침(Data Protection Directive)을 전면 수정하여 2012년 데이터 보호 개혁(Data Protection Reform)을 실시하고 ‘데이터 보호법(Data Protection Regulation)’ 개정안을 발표하였다(한은영, 2014).

데이터 보호법 개정안의 주요 내용을 살펴보면 회원국 간 상이한 법 적용으로 인한 혼란을 제거하고자 지침(directive)이 아닌 상위규범인 일반규정(regulation)으로 체계를 변경하였고, 회원국 국민의 개인정보를 보호하기 위해 법률의 효력을 글로벌 기업에게까지 확장하였다. 또한 기존보다 명확하고 세분화된 개인의 동의를 이루어지도록 명시하고 있으며, 새로운 개인정보보호 측면에서 ‘잊혀질 권리(right to be forgotten)’에 대한 내용을 추가하였다(김상우, 2012).

2012년 개정안에 명시된 잊혀질 권리는 ‘삭제할 권리(right to be

forgotten and to erase)’와 함께 정보 주체의 자기정보통제권을 보호하는 규정이다. 정보의 수집·처리 목적에 부합하지 않거나 정보 처리에 대한 동의를 철회하거나 동의의 기간이 만료되었을 경우, 정보 주체가 정보처리를 반대하는 등의 경우에 정보의 삭제나 폐기, 오류에 대한 정정을 요구할 수 있는 강력한 권리이다(정영화, 2012).

이처럼 개인정보보호에 대한 강력한 규제를 도입하고 국가의 역할을 강조하는 가운데, 2014년 7월 유럽위원회는 데이터주도 경제(Data-Driven Economies)를 추진하기 위한 실행계획인, Communication을 발표하였다. 2012년부터 빅데이터 산업에 대한 관심을 표명한 유럽위원회는 실행계획을 통해 공공부문에서 선도적으로 신뢰할 만한 데이터를 공개하고, 유럽 연합 회원국의 네트워크를 강화하며, 데이터셋을 활용할 수 있는 전문 인력을 양성하데 집중할 것임을 밝혔다(한국방송통신전파진흥원, 2014).

또한 개인정보보호에 강력한 중점을 둔 유럽연합이 빅데이터 산업 양성과 개인정보보호의 균형을 맞추기 위해 익명화(data anonymisation)와 필명화(pseudonymisation), 데이터 최소화, 개인정보 위험분석 등에 관한 가이드를 마련하여 기업 등에 제공할 예정이다(한국방송통신전파진흥원, 2014). 유럽연합은 정보 주체의 프라이버시 보호하기 위해 개인정보 자기결정권을 보장하고 개인정보처리자의 책임과 의무사항을 확대하는데 중점을 두는 동시에, 빅데이터 산업의 활성화를 추진하기 위한 방안을 모색하고 있다.

(3) 일본

일본은 2005년 4월 개인정보보호에 관한 법률을 시행하였다. 개인정보 보호에 관한 법률에서는 개인정보취급자의 의무로 개인정보보호지침, 목적 외 이용금지, 부정취득 금지, 안전관리조치 마련, 목적 외 이용 및 제3자 제공 시 동의 획득 필요, 본인이 동의한 방법으로 공시 등이 있다. 정보 주체는 개인정보취급자에게 대한 통지 또는 공표권, 공개요청권, 정

정요청권, 이용정지권을 행사할 수 있다(전은정 외, 2012b).

최근 일본에서 빅데이터 활용을 위해 개인정보보호법 개정을 추진하고 있는데, 10년 만에 이루어지는 개정을 앞두고 발표된 초안의 추진 과정은 다음과 같다. 정부 산하의 고도정보통신네트워크사회추진 전략본부(高度情報通信ネットワーク 社會推進戰略本部, 이하 IT 종합전략본부)는 2013년 12월 ‘개인 데이터의 활용에 관한 제도 재검토(パーソナルデータの利活用に關する制度見直し方針)’ 정책을 발표하였다. 이에 의거하여 ‘개인 데이터에 관한 검토회(パーソナルデータに關する檢討會, 이하 검토회)’를 설치하고 빅데이터 시대에서 산업의 혁신을 위해 개인 데이터의 활용을 촉진하는 활동 목표를 상정하였다. 이와 함께 개인정보보호와 관련된 신속한 대응 체계를 구축하는 등 개인정보보호를 위한 방안을 마련하고자 하였다(한국인터넷진흥원, 2014).

그 결과 11차례의 회의를 통해 2014년 6월 개인정보보호법 개정안을 마련하였다. 개정 초안이 중점적으로 다루고 있는 주제는 먼저, 본인의 동의가 없어도 데이터를 활용할 수 있도록 하는 기틀을 도입하는 것이다. 또한 개인정보보호를 위해 민간 자율 규제를 유도하고 개인정보보호 감독을 위한 제3의 감독기관을 설치하는 것이다(한국인터넷진흥원, 2014).

개정안의 주요 내용을 살펴보면 첫째, 개인을 식별할 수 있는 가능성을 제거하는 가공과정 이후에는 개인정보를 활용하고 제3자에게 제공할 때 동의가 필요한 데이터일지라도 자유롭게 활용하도록 하고 있다. 이때 데이터를 가공하는 방법에 대해서는 자율 규제를 원칙으로 하되, 제3의 감독기관에서 인증을 받도록 조치하였다(한국인터넷진흥원, 2014).

둘째, 개인정보보호의 대상과 취급 방법을 구체화하고 민간 자율 규제를 유도하고 있다. 개인정보보호 대상에는 지문, 얼굴 데이터 등 신체 특징과 관련된 데이터를 포함하여 개인정보를 명확히 하였다, 그 외 데이터의 경우 제3의 감독기관에 따라 보호 여부가 결정된다. 검토회는 정부, 기업, 소비자, 관련 전문가들과 함께 개인 데이터의 활용과 개인정보보호를 양립할 수 있는 자율 규제 원칙을 수립하고자 한다. 규칙을 담당하는

자율단체는 사업자에게 필요한 조치를 취하고, 제3의 감독기관의 규제를 준수하도록 명시하였다(한국인터넷진흥원, 2014).

세 번째로 제3자 감독기관의 설치를 명문화하고 그 권한과 역할을 규정하고 있다. 감독기관을 통해 전문적 지식을 갖춘 집단이 신속하고 적절하게 관련법을 집행하고, 개인 데이터의 보호와 활용 간 균형을 이루고자 한다. 감독기관에 개인정보를 취급하는 사업자에 대해 여러 권한을 부여하고 있다. 먼저, 개인정보를 취급하는 사업자에 대한 현장 감사의 권한을 지니며, 현재 내각 총리가 수행하고 있는 개인정보보호 단체 인증 및 인증 취소 관련 기능을 수행한다. 또한 민간 주도의 개인정보보호 자율 규제를 수립하는데 해당 규제를 인증하고 규제를 운영하는 단체에 대한 감독을 실시하는 역할을 담당한다(한국인터넷진흥원, 2014).

3) 빅데이터 개인정보보호 가이드라인

일본과 비슷한 시기인 2013년 12월 18일, 미래창조과학부와 방송통신위원회는 ‘빅데이터 개인정보보호 가이드라인(안)’을 발표하였다. 그 후 다양한 이해관계가 존재하고 기존 법규들과 충돌하는 문제로 인해 가이드라인은 논란이 되었고 수정요구를 받아왔다. 가이드라인 초안이 발표된 이후 총 4차례의 공개 토론회 등 사회적 논의를 거쳐 2014년 11월 20일에 4차 수정안이, 그 후 12월 23일에 수정안이 다시 발표되었다. 2013년 초안과 2014년 가이드라인 수정안의 주요 내용은 [표 2-3]과 같다.

초안과 달리 2014년 발표된 가이드라인에서 주목할 점은 ‘비식별화 조치’를 강조하고 있는 것이다. 비식별화란 “데이터 값 삭제, 가명처리, 총계처리, 범주화, 데이터 마스킹 등을 통해 개인정보의 일부 또는 전부를 삭제하거나 대체함으로써 다른 정보와 쉽게 결합하여도 특정 개인을 식별할 수 없도록 하는 조치”이다(제2조). 정보통신서비스 제공자가 개인정보가 포함된 공개된 정보와 이용내역정보를 수집할 때는 비식별화 조치가 선행되어야 하며 그 후에 수집·저장·조합·분석 및 제3자 제공 등이

가능하도록 명시하고 있다(제3조, 제4조, 제5조, 제10조).

또한 빅데이터 처리 사실과 목적 등을 공개하여 투명성을 확보하도록 하고 있다(제4조, 제5조, 제9조). 비식별화 조치를 취한 후 빅데이터 처리 사실과 목적을 공개하고, ‘정보 활용 거부 페이지 링크’를 제공하여 이용자가 정보 활용 거부권을 행사할 수 있도록 한 것이다. 또한 이용자 이외의 자로부터 수집한 개인정보를 처리할 때 이용자가 수집 출처와 목적, 개인정보 처리 정지 요구권을 행사할 수 있음을 공개하도록 하였다.

빅데이터가 처리되는 과정에서 혹은 생성정보에서 개인정보가 재식별되었을 경우에 그 정보를 즉시 파기하거나 추가적으로 비식별화 조치를 하도록 명시하고 있다(제3조, 제6조). 비식별화 조치가 취해진 정보를 수집하여 정보 처리시스템에 저장·관리할 때 접근 통제장치를 운영하거나 백신 소프트웨어를 설치하는 등의 기술적·관리적 보호조치를 취하도록 하고 있다(제3조 제2항). 이 외에도 개인의 사상·신념, 정치적 견해와 같은 민감정보를 생성하기 위한 목적으로 정보를 수집, 이용, 분석 등 처리하는 것을 금지하고 있다.

2013년 초안에서 논란이 되었던 ‘별도로 정보주체의 동의를 얻지 않아도’ 정보를 수집·이용할 수 있다고 명시한 부분을 수정하였다. 그 결과 2014년 가이드라인에서는 ‘비식별화 조치’가 선행되거나 ‘이용자의 동의’가 있을 때 정보의 수집과 활용이 가능하도록 하였다. 하지만 여전히 빅데이터 산업의 특성상 사업자의 정보 활용에 대해 정보주체에게 사전동의를 받기가 어렵다는 점을 강조하며 ‘정보활용 거부권’을 대안으로 제시하고 있다.

[표 2-3] 빅데이터 개인정보보호 가이드라인(안)의 주요내용

조항	가이드라인 주요내용	
	2013년 초안	2014년 수정안
정의	[제2조 제1항] ‘공개된 개인정보’란 정보주체 및 정당한 권한이 있는 자에 의해 제한 없이 일반 공중에 공개된 부호·문자·음성·음향 등의 정보로서 생존하는 개인을 식별할 수 있는 정보 및 다른 정보와 쉽게 결합하여 개인을 식별할 수 있는 정보를 말한다.	[제2조 제1항] ‘공개된 정보’란 이용자 및 정당한 권한이 있는 자에 의해 공개 대상이나 목적의 제한 없이 합법적으로 일반 공중에 공개된 부호·문자·음성·음향 및 영상 등의 정보를 말한다.
개인정보의 보호	-	[제3조 제1항] 개인정보가 포함된 공개된 정보 및 이용내역정보는 <u>비식별화 조치를 취한 후</u> 수집·저장·조합·분석 등 처리하여야 한다. 비식별화 조치된 공개된 정보 및 이용내역정보를 조합·분석 등 처리하는 과정에서 개인정보가 생성되지 않도록 하여야 한다. 다만, 개인정보가 생성되는 경우에는 지체없이 파기하거나 비식별화 조치를 취하여야 한다.
공개된 정보의 수집·이용	[제3조 제1항] 정보통신서비스 제공자가 공개된 개인정보를 수집하고자 하는 경우에는 <u>별도로 정부주체의 동의를 얻지 아니하여도 된다.</u>	[제4조 제1항] 정보통신서비스 제공자가 개인정보가 포함된 공개된 정보를 <u>비식별화 조치한 경우에는 이용자의 동의 없이 수집·이용할 수 있다.</u> 다만, 이용자의 동의를 받거나 법령상 허용하는 경우에는 비식별화 조치를 취하지 아니하고 수집·이용할 수 있다.

[표 2-3] 빅데이터 개인정보보호 가이드라인(안)의 주요내용 계속

조항	가이드라인 주요내용	
	2013년 12월	2014년 12월
이용내역 정보의 수집·이용	[제4조 제1항] 정보통신서비스 제공자가 정보통신서비스 제공과 관련하여 조합, 분석, 또는 처리를 목적으로 이용내역을 수집하는 경우 별도로 정보주체의 동의를 얻지 아니하여도 된다.	[제5조 1항] 정보통신서비스 제공자가 다음 각 호의 경우를 제외하고는 이용자의 동의를 받거나 <u>비식별화 조치를 취한 후</u> 이용내역정보를 수집·이용할 수 있다.
새로운 정보의 생성	[제5조 제1항] 정보통신서비스 제공자는 정보조합·분석·처리 시스템을 통해 공개된 개인정보 및 이용내역정보 등을 활용하여 새로운 개인정보를 생성하는 경우 별도로 정보주체의 동의를 얻지 아니하여도 된다.	[제6조 제1항] 정보통신서비스 제공자는 <u>비식별화 조치</u> 하여 수집한 공개된 정보 및 이용내역정보를 정보 처리시스템을 통해 조합·분석하여 새로운 정보를 생성할 수 있다. 다만, 새롭게 생성된 정보에 개인정보가 포함되어 있을 경우, 즉시 파기하거나 비식별화 조치를 취하여야 한다.
공개된 정보 및 이용내역 정보의 이용	[제10조 제1항] 정보통신서비스 제공자는 공개된 개인정보 등을 정보주체의 별도 동의 없이 자신의 서비스 제공업무 수행을 위해 내부에서 이용할 수 있다. 다만, 정보주체가 거부 의사를 표시한 때에는 그러하지 아니하다.	[제9조 제1항] 정보통신서비스 제공자는 비식별화 처리된 공개된 정보 및 이용내역정보를 자신의 서비스 제공업무 수행을 위해 내부에서 이용할 수 있다. 다만, 이용자가 거부 의사를 표시한 때에는 그러하지 아니하다.
개인정보의 제3자 제공	[제11조 제1항] 정보통신서비스 제공자는 공개된 개인정보 등을 제3자에게 제공하려면 제공받는 자, 이용 목적, 항목, 보유 및 이용기간 등을 알리고 사전 동의를 받아야 한다. 다만, <u>공개된 개인정보</u> 는 정보주체의 동의없이 제3자에게 제공할 수 있다.	[제10조] 정보통신서비스 제공자는 개인정보가 포함된 공개된 정보, 이용내역정보, 생성 정보의 경우, <u>이용자의 동의를 얻어 제3자에게 제공할 수 있다.</u> 다만, 비식별화 처리된 공개된 정보, 이용내역정보, 생성 정보는 이용자 동의 없이 제3자 제공이 가능하다.

제 2 절 빅데이터 환경에서의 개인정보 활용에 관한 논의

1. 정보주체의 권리에 관한 논의

2013년 빅데이터 개인정보보호 가이드라인 초안이 발표되고 현재까지 논란이 되고 있는 쟁점은 빅데이터 개인정보에 대한 정보주체의 ‘자기결정권’과 관련이 있다. 김선남과 이환수에 의하면(2014) 가이드라인은 공개된 개인정보에 대해서 정부주체 및 정보에 대한 권한이 있는 주체가 정보를 공개했다는 사실만으로 정보처리자의 수집 및 이용 등의 행위를 정당화하고 있다. 가이드라인에 명시된 바를 살펴보면, 제3조 제1항은 정보통신서비스 제공자가 공개된 개인정보를 수집하고자 하는 경우에는 별도로 정부주체의 동의를 얻지 않아도 됨을 보여준다. 이는 개인정보 자기결정권을 무력화하는 것으로, 이는 제5조와 제10조, 제11조에 공개된 개인정보의 분석, 이용, 제3자 제공에서 정부주체의 동의가 필요하지 않다고 명시한 바에서도 확인할 수 있다.

이용내역정보에 있어서도 가이드라인 제4조에서는 정보주체로부터 별도의 동의를 얻지 않아도 정보처리자가 수집 및 이용할 수 있도록 하고 있다. 명시된 방식은 온라인 사이트에서 생성된 소비자의 위치정보, 검색 정보 등을 사업자가 수집하여 개인의 민감한 사생활 정보를 파악할 수 있는 문제를 내포하고 있다. 또한 이용내역정보를 사업자가 수집하는 것을 차단하기 위해서는 거부 의사를 표시해야 하는(Opt-out) 방식을 도입하고 있다(김선남·이환수, 2014). 이용내역정보가 발생하는 시점마다 동의를 받기 어렵고, 정보통신서비스와 직·간접적으로 관련되어 있어 정보를 수집한다는 것에 대한 정보 주체의 예측이 가능하다는 것에 기인한 것이다.

옵트아웃 방식을 보완하기 위해 제4조 제2항에서는 이용내역정보를 수집하는 경우 정보가 수집·활용 된다는 사실과 목적을 정보 주체가 언제든지 쉽게 확인할 수 있도록 공개하는 것을 명시하였다. 표면적으로 정보주체의 개인정보 자기결정권을 보장하기 위한 것으로 보이지만, 실질

적으로 이러한 명시를 통해 소비자가 자신의 정보가 어떠한 방향으로 활용되고 있는지 확인하기 어려워 실효성이 없다는 지적을 받아 왔다(김선남·이환수, 2014).

과생정보의 측면에서도 온라인상 공개된 정보와 기존의 정보의 결합을 통해 개인의 사상과 신념과 더불어 개인의 건강 및 성생활 등에 관한 민감한 사생활을 침해할 우려가 존재한다(김승한, 2013). 또한 페이스북과 같은 SNS에 작성된 개인의 견해와 다른 정보의 결합을 통해 활용될 문제가 존재한다. 하지만 정보주체가 SNS를 통해 일상 정보를 게재한 목적이 빅데이터 분석과 같은 다른 목적으로 활용되기를 원해서가 아니라, 타인과 관계를 형성하고 공동체에 대한 동질감 및 신뢰를 위한 것이다(이규철, 2013). 따라서 정보주체의 동의 없이 과생정보가 사업자에 의해 활용된다면 개인정보 자기결정권을 침해하는 것이다.

정보처리자가 다양한 정보를 수집·분석하여 식별 가능성이 존재하는 새로운 개인정보를 취득할 수 있게 되고 이때 정보주체의 개인정보 자기결정권이 침해될 위험이 높다. 문제를 해결하기 위해 개인정보보호법 내 개인정보 정정 또는 삭제에 관한 권리를 요구할 수 있는 정보주체의 권리를 구체적으로 명시하는 방안이 추가되는 방안이 제시되었다(김선남·이환수, 2014).

2014년 발표된 수정된 가이드라인은 정보주체의 개인정보 자기결정권을 보장하는 방안으로 제한된 ‘동의권’ 및 초안과 같은 ‘정보 활용 거부권’을 제시하였다. 그럼에도 불구하고 수정안 제4조를 보면 공개된 정보가 ‘비식별화’ 되었을 경우 사업자는 여전히 정보주체의 동의 없이 정보를 활용할 수 있다. 제5조에서 이용내역정보를 활용하기 위해서 동의를 받아야 한다고 명시하고 있다 하지만 ‘정보통신서비스의 제공에 관한 계약을 이행하기 위하여 필요한 이용내역정보로서 경제적·기술적인 사유로 통상적인 동의를 받는 것이 뚜렷하게 곤란한 경우’ 등과 같이 사실상 동의권이 보장되지 않는 경우가 빈번하게 발생할 가능성이 존재한다. 또한 새로운 정보를 생성하는 경우에 정보처리자가 공개된 정보 및 이용내역정보를 비식별화하였다면 여전히 정보 주체의 동의 없이 정보가 활용될

수 있다.

정보처리자가 비식별화 조치를 취한 공개된 정보 및 이용내역정보를 활용할 때 그 사실과 목적을 이용자가 쉽게 확인할 수 있도록 공개하도록 명시하고 있다. 또한 개인정보 취급방침을 통해 정보의 활용을 거부할 수 있는 방법을 투명하게 공개하도록 하고, ‘정보 활용 거부 페이지 링크’를 제공하도록 명시하였다. 하지만 가이드라인 초안에서 지적된 바와 같이 명시된 조치가 실효성을 발휘하기 위한 명확한 방침이 제시되고 있지 않았다.

2. 정보처리자의 의무에 관한 논의

정보처리자에게 정보주체로부터 사전동의를 받아야 하는 의무를 축소해주는 대신 빅데이터 개인정보를 보호하기 위한 조치로 ‘비식별화’와 ‘기술적·관리적 보호조치’를 의무화하는 가이드라인의 구체적인 내용은 다음과 같다.

비식별화는 ‘데이터 값 삭제, 가명처리, 총계처리, 범주화, 데이터 마스킹 등을 통해 개인정보의 일부 또는 전부를 삭제하거나 대체함으로써 다른 정보와 쉽게 결합하여도 특정 개인을 식별할 수 없도록 하는 조치’이다(제2조 제4항). 가이드라인에서는 비식별화 조치에 대한 규정만을 제시하고 있으며 구체적인 조치를 위한 수단 및 방법 등은 사업자에게 일임하였다. 정보주체가 자신의 개인정보가 비식별화 된다는 사실을 인지하더라도 단순히 비식별화 조치라는 전제만으로 개인정보의 활용에 신뢰하고 동의할지 명확히 알 수 없다.

공개된 정보를 수집·저장하고 관리할 때 정보처리자가 시행하는 기술적·관리적 보호조치는 제3조 제2항에 명시되어 있다. 구체적으로 정보처리자는 1) 불법적인 접근을 차단하기 위한 침입차단시스템 등 접근 통제장치를 설치하고 운영하며, 2) 접속 기록의 위·변조를 방지하기 위한 조치와 3) 백신 소프트웨어 설치·운영 등 악성프로그램에 의한 침해를 방

지하기 위한 조치를 취해야 한다. 개인정보보호위원회에 따르면 2008년 이래로 발생한 총 82건의 유출 건수 중 80%에 이르는 65건이 해킹이나 내부직원·외부용역업체 직원의 고의적인 유출에 의한 것이다. 해킹으로 인한 공격에 대응하지 못한 것도 문제지만 내부 직원 등에 의한 유출을 방지하지 못한 관리의 책임 또한 존재한다. 이와 같이 기존에 관리하던 개인정보도 제대로 보호되지 않고 있는 상황에서 기업들에게 더욱 방대한 양의 정보를 수집·활용할 수 있도록 허가하고, 이를 제대로 보호할 것이라고 기대하기란 매우 어렵다.

일각에서는 빅데이터 산업 활성화를 통해 새로운 일자리를 창출하고, 경제적 부가가치를 생산해낼 수 있음을 강조하여 지나친 정보보호를 강조하는 것이 걸림돌이 될 수 있다고 주장한다. 따라서 정보보호의 규율이 약화되는 점을 보완하기 위해 강력한 소비자 구제책으로써 징벌적 손해배상제도를 도입하거나 집단소송제도를 보완하는 방안을 제시하기도 한다(정원준, 2014). 일반적으로 개인정보 유출로 인한 소송으로 1인당 배상을 받는 금액은 10만원에서 30만원에 불과한데 비해 소송은 수년 간 지속되기 때문에 소비자들에게 불리하다. 이를 보완하기 위해 집단소송제도와 징벌적 손해배상제도가 개인정보 유출 소송에 적용된다면 수많은 소비자들이 피해를 보상받을 수 있으며, 기업에게 개인정보 보호의 의무에 대한 책임감을 각인시킬 수 있을 것이다. 하지만 빅데이터 환경에서 생산되는 개인정보의 유출에 대한 명확한 책임소재를 규정하고, 앞서 언급한 강력한 제도를 개인정보보호에 도입하기 위해서는 오랜 기간에 걸친 사회적 논의가 필요하므로 근시일내에 이루어질 가능성이 희박하다.

제 3 절 개인정보에 대한 소비자인식 관련 연구

1. 개인정보 침해에 대한 인식

기존 개인정보와 관련된 연구에서는 주로 개인정보가 유출되어 발생할 위험에 대한 소비자들의 우려와 지각을 다루었다. 먼저 개인정보에 대한 소비자인식 중 개인정보 침해에 대해 우려(concern)와 관련된 변수를 측정하는 연구는 다음과 같다. 프라이버시 정책 운영과 관련하여 소비자들의 우려에 영향을 미치는 하위개념을 측정하였다. 그 중 사전동의 없이 개인정보가 2차적으로 사용되는 것(secondary use)과 개인정보에 부적절한 접근(improper access)이 이루어지는 것을 가장 우려하였다(Smith et. al., 1996).

개인정보의 식별가능성에 따른 우려의 차이가 존재하며 개인 식별정보를 제공하는 것에 대해 우려가 더 높다는 사실이 밝혀졌다. 이와 달리 이용자가 비식별정보는 보다 쉽게 제공하는 경향이 있으며, 인터넷이 제공하는 편익에 대한 인식에 비해 프라이버시에 대한 우려가 더 높다는 사실이 발견되었다(Cranor et. al., 1999),

개인정보가 유출될 경우에 발생할 피해에 대한 소비자의 위험지각(perceived risk)을 측정한 연구도 존재한다. 김형길과 김정희(2000)는 전자상거래에서 정보유출로 인해 사생활이 침해될 위험을 소비자가 지각하고 있음을 밝혀내었다. 또한 김소라와 이기춘(2006) 연구를 살펴보면 소비자가 지각하는 개인정보 유출위험이 전반적으로 높았으며, 개인정보별 위험지각에 차이가 있었다.

또한 전상민(2015)은 개인정보 유출에 대한 위험지각이 부정적 감정에 영향을 준다는 사실을 밝혀내었다. 불쾌감과 각성, 무력감과 같은 부정적 감정이 개인정보유출에 대한 지각된 위험에 영향을 주고, 부정적 감정이 부정적 행동에 유의한 영향을 미쳤다.

2. 개인정보 및 개인정보 제공에 대한 인식

지금까지는 개인정보 침해로부터 정보를 보호하기 위해 마련된 관련법을 고찰한 연구가 대부분이었다. 하지만 정작 정보 주체인 소비자의 입장에서 개인정보란 무엇인지, 사업자가 개인정보를 활용하는 것을 어떻게 인식하고 있는지에 대한 연구는 부족한 실정이다. 비교적 최근 들어 등장한 소비자의 입장에서 개인정보에 대한 인식을 연구한 논문들은 다음과 같다.

먼저 김민성과 김성태(2014)는 이용자의 측면에서 개인정보 유형을 분류하고자 기존에 정리된 개인정보들을 제시하고 이를 개인정보라고 인식하는 정도를 살펴보았다. 이름, 나이, 생년월일에서 인터넷 홈페이지 등의 방문내역, 아이디나 비밀번호, 사진·동영상·음악파일에 이르기까지 매우 광범위한 정보를 사용자에게 인식에 따라 전용 신원 정보, 온라인 기록 정보, 라이프스타일 정보, 통용 인적 정보 네 가지로 분류하였다.

김지온과 박지경(2014)은 개인정보 항목을 제시하고 이에 대해 개인정보라고 인지(recognition)하는 정도를 측정하였다. 이름, 주민등록번호/계좌번호, 어떤 개인만의 특이사항, 주소, 자동차번호 등 개인정보에 대한 성별과 학제별, 개인정보보호법 인지 여부별 차이를 서술하였다. 전반적으로 이름과 주민등록번호/계좌번호를 개인정보로 인지하는 수준이 높았으며 자동차번호 및 어떤 개인만의 특이사항은 인지수준이 낮았다.

개인정보 자체에 대한 인식 외에도 개인정보를 사업자에게 제공하는 것에 대한 소비자의 인식을 살펴본 연구도 이루어졌다. 구혜경과 나종연(2014)은 개인정보를 제3자에게 제공하는 것에 대한 소비자의 인식(perception)을 연구하였다. 소비자들은 서비스 제공에 필수적으로 필요한 상황이 아닐 경우 자신의 개인정보가 공유되는 것을 부정적으로 인식하고 있었다. 또한 소비자들은 사업자가 개인정보를 제3자에게 제공하는 경우 이에 대해 명확히 알리고 쉽게 제3자에 대해 확인할 수 있어야 하며, 이에 대한 동의를 중요시 여기고 있음을 밝혀내었다.

제 3 장 연구문제 및 방법

제 1 절 연구문제

본 연구는 빅데이터 환경에서의 개인정보에 대한 인식 수준은 어떠한가
며 소비자집단에 따라 차이가 있는지 살펴보고자 한다. 또한 개인정보
활용에 대한 인식 수준과 이에 영향을 미치는 변수에 대해 알아보고 결
과적으로 빅데이터 환경에서의 개인정보에 대한 소비자의 인식을 심층적
으로 이해하고자 한다.

[연구문제 1] 빅데이터 환경에서의 개인정보에 대한 소비자의 인식수준
은 어떠한가?

1-1. 빅데이터 환경에서의 정보를 자신의 개인정보로 인식하는
수준은 어떠한가?

1-2. 개인정보 관련 행동, 경험 및 기타 인구통계학적 변수에 따
라 개인정보에 대한 인식수준은 어떠한 차이를 보이는가?

[연구문제 2] 빅데이터 환경에서의 개인정보 활용에 대한 소비자의 인식
수준은 어떠한가?

2-1. 빅데이터 환경에서의 개인정보 활용에 대한 인식수준은 어
떠한가?

2-2. 빅데이터 환경에서의 개인정보 활용에 대한 동의정도에 영
향을 미치는 요인은 무엇인가?

제 2 절 연구방법

1. 조사대상 및 자료수집 방법

빅데이터 환경에서의 개인정보와 그 활용에 대한 인식수준, 그리고 영향요인을 밝히기 위해 본 연구는 조사전문기관인 엠브레인의 온라인패널을 대상으로 설문조사를 실시하여 자료를 수집하였다. 본 연구는 한국인터넷진흥원에서 발표한 2014년 인터넷이용실태조사 결과를 바탕으로 인터넷 이용률이 높은 10대에서 50대의 소비자를 대상으로 연령과 성별로 할당하였다. 서울대학교 생명윤리심의위원회(IRB)의 심의를 받은 내용 및 설문지를 통해 진행되었다.

본 조사를 수행하기에 앞서 2015년 4월 27일부터 28일까지 이틀에 걸쳐 예비조사를 실시하였다. 총 40명을 대상으로 예비조사를 실시한 결과를 바탕으로 문항의 신뢰도 및 적합성을 검토하였다. 본 조사는 2015년 4월 29일부터 5월 6일까지 진행되었으며, 총 580부의 자료 중 연구에 적합하지 않은 설문을 제외하고 최종적으로 574부를 분석 대상으로 사용하였다.

2. 변수의 정의 및 측정도구

1) 개인정보 관련 행동, 경험 변수

(1) 개인정보 보호행동

온라인 개인정보 보호행동은 ‘소비자들이 자신의 개인정보를 보호하기 위해 온라인 상에서 행하는 행동’을 말한다. 개인정보 보호행동의 수준을 측정하기 위해 구체적으로 적극적인 개인정보 공개 거부와 관련된 행동과 상대적으로 소극적인 개인정보관련 정책이나 사안들에 관심을 갖는 행동으로 구성하였다. 김정은(2008)은 Youn(2005), Metzger(2006) 등의 연구를 참고하여 개인정보 보호행동 측정 척도를 총 15개 문항으로 구성하였다.

본 연구는 김정은(2008)의 설문문항을 수정하여 온라인 상 개인정보 보호행동을 측정하는 총 4개 문항을 구성하였다. 개인정보 공개 거부와 관련된 행동으로 ‘인터넷 사이트 가입 시 너무 많은 개인정보를 요구하면 다른 사이트로 간다’와 ‘인터넷 사이트 가입 시 불필요한 정보는 적지 않는다’ 등 두 가지 문항을 제시하였다. 개인정보 관련 정책이나 사안에 관심을 갖는 행동은 ‘개인정보를 보호하기 위한 방법을 찾아본다’와 ‘인터넷 사이트에 가입 시 사이트 내 개인정보보호정책이 있는지 확인하다’ 등 두 가지 문항으로 측정하였다.

(2) 개인정보 약관 확인행동

온라인 사이트에 가입할 때 사업자의 개인정보 수집·이용과 관련된 약관을 확인하는지 행동을 측정하기 위해 나종연(2013) 연구를 참고하였다. 실제로 온라인상에서 개인정보 수집·이용 약관에 동의여부를 클릭하거나 체크해야 하는 경우의 행동을 묘사하여 제시하였다. 구체적으로 ‘내용을 읽지 않고 체크해야 하는 부분만 찾아서 체크하기’, ‘대충 스크롤을 내려 보고 체크하기’, ‘눈에 띄는 부분만 확인하고 체크하기’ 그리고 ‘꼼꼼히 읽어보고 체크하기’ 등 네 가지 문항으로 구성하였다.

(3) 개인정보 유출 경험

지난 1년 동안 주민등록번호나 이름, 주소 등의 개인정보가 유출된 경험이 있는지를 측정하였다. 한국인터넷진흥원에서 발표한 2014년 정보보호 실태조사(개인 편)를 참고하였다. 예비조사를 통해 응답자들이 자신의 개인정보가 유출된 것을 확신하지 못하고 있거나 명확하게 정보가 유출된 횟수를 기억하지 못한다는 것을 고려하여 문항을 재구성하였다. 기존에 개인정보 유출 경험의 횟수를 제시하였던 문항을 수정하여 '개인정보 유출 경험이 있다, 없다, 모른다'로 구성하였다.

2) 빅데이터 환경에서의 개인정보에 대한 인식

본 연구에서는 빅데이터 환경에서의 개인정보에 대한 인식을 '빅데이터 환경에서 발생하는 정보를 자신의 개인정보라고 생각하는 정도'라고 정의하였다. 기존과는 달리 온라인 사이트에 가입할 당시에 제공하는 정보 외에도 사이트 가입 후 소비자들이 적극적으로 생산하는 정보의 양이 폭발적으로 증가하였다. 새로운 정보들에 대한 소비자들의 근본적인 인식, 즉 빅데이터 환경에서 발생하는 정보를 자신의 개인정보라고 생각하는지 살펴보았다.

빅데이터 환경에서 생성되는 개인정보를 2014년 개인정보보호 가이드라인에서 제시하고 있는 공개된 정보와 이용내역정보 두 영역으로 나누었다. 또한 공개된 정보는 정보 공개 주체를 자신과 타인으로 나누어 제시하였다. 구체적으로 공개된 정보에 속한 개별 정보들은 페이스북 등 소셜네트워크서비스에서 소비자가 전체 공개할 수 있는 정보들로 구성하였다. 이용내역정보에는 가이드라인에 예시로 제시된 정보와 함께 나중연 등(2012)의 연구에서 분류한 인터넷 행태에 관한 정보인 클릭한 광고, 검색어, 방문한 사이트를 포함하였다. 공개된 정보 10개, 이용내역정보 5개, 총 15개의 개인정보는 [표 3-1]과 같다.

구체적으로 ‘나는 (자신이 전체 공개한 일반 인적 사항, 온라인 인적 사항, 글, 사진 및 동영상, 위치정보, 사회적 정보, 타인이 전체 공개한 내가 포함된 글, 내가 포함된 사진 및 동영상, 내가 포함된 위치정보, 내가 포함된 사회적 정보, 사이트 접속기록, 사이트 접속 실시간 기록, 사이트 클릭기록, 사이트 검색기록, 사이트 거래기록)을 나의 개인정보라고 생각한다’라는 질문에 대해 5점 리커트 척도(전혀 그렇지 않다-매우 그렇다)로 측정하였다.

[표 3-1] 빅데이터 환경에서의 개인정보 유형과 종류

유형		종류
공개된 정보	자신	일반 인적 사항 (이름, 성별, 학력, 직업 등)
		온라인 인적 사항 (아이디, 이메일 등)
		글
		사진 및 동영상
		위치정보
	사회적 정보 (SNS 인맥, 동호회 등)	
	타인	내가 포함된 글
		내가 포함된 사진 및 동영상
		내가 포함된 위치정보
		내가 포함된 사회적 정보 (SNS 인맥, 동호회 등)
이용내역정보	사이트 접속기록 (접속 시간, 장소 등)	
	사이트 접속 실시간 기록 (현재 위치 정보 등)	
	사이트 클릭기록 (광고, 방문 사이트 등)	
	사이트 검색기록 (검색어 등)	
	사이트 거래기록 (쇼핑 내역 등)	

3) 빅데이터 환경에서의 개인정보 활용에 대한 인식

(1) 개인정보 활용에서의 의사 반영에 대한 인식

빅데이터 환경에서 개인정보가 사업자에 의해 활용될 때 소비자의 의사가 어느 정도 반영될 것이라고 생각하는지를 측정하였다. 구체적으로 개인정보가 수집·저장·조합·분석 등 처리되는 과정에 또한 개인정보가 처리되는 범위에 소비자 자신의 의사가 반영될 것이라고 생각하는지를 5점 리커트 척도로 측정하였다.

(2) 빅데이터 환경에서의 개인정보 활용에 동의하는 정도

빅데이터 환경에서 사업자가 개인정보를 활용하는 것에 소비자가 어느 정도 동의하지 측정하였다. 빅데이터 환경에서의 개인정보를 두 가지 영역, 즉 공개된 정보와 이용내역정보로 구분하여 설문하였다. 그리고 각 정보에 있어서 사업자가 개인정보를 활용할 때 취하는 조치의 강도를 빅데이터 개인정보보호 가이드라인을 바탕으로 구성하였다. 구체적으로 가이드라인에서 제시한 개인정보에 대한 비식별화 조치를 상기시키기 위해 “사업자가 나의 개인정보를 모두 내가 누군지 알 수 없게 처리한 상태입니다.”라는 설명 문구를 문항 앞에 제시하였다. 그 후 사업자의 조치에 따라 총 네 가지 경우를 제시하여 사업자의 개인정보 활용에 대한 소비자의 동의정도를 질문하였다. 소극적 조치는 ‘별다른 조치를 취하지 않았을 경우’와 ‘활용에 대해 쉽게 확인할 수 있도록 공개하였을 경우’ 등 두 문항으로 구성되어 있으며, 적극적 조치는 ‘소비자에게 고지하였을 경우’와 ‘소비자의 동의를 받았을 경우’ 등 두 문항이다. 두 가지 정보 범주에 대해 네 가지 사업자의 조치를 제시하여 측정한 8개의 문항은 [표 3-2]와 같다.

[표 3-2] 개인정보 활용에 대한 동의정도 측정 문항

구분	문항
공개된 정보	내가 모두에게 공개한 것이므로 별다른 조치없이 사업자가 활용해도 된다.
	사업자가 활용한다는 것을 쉽게 확인할 수 있으면, 사업자가 활용해도 된다.
	사업자가 활용할 때 나에게 알리면 활용해도 된다.
	사업자가 활용할 때 나에게 동의를 받는다면 활용해도 된다.
이용내역정보	온라인 사이트에 가입하는 것만으로도 사업자가 활용해도 된다.
	사업자가 활용한다는 것을 쉽게 확인할 수 있으면, 사업자가 활용해도 된다.
	사업자가 활용할 때 나에게 알리면 활용해도 된다.
	사업자가 활용할 때 나에게 동의를 받는다면 활용해도 된다.

또한 사업자가 빅데이터 환경에서 생산되는 개인정보를 활용하여 행하는 활동 범위, 즉 혜택의 제공에 따라 소비자의 동의정도가 변화하는 지 측정하였다. 개인정보를 활용하여 일반적 광고, 맞춤형 광고, 제품 개발 및 홍보활동을 하는 것 등 사업자의 세 가지 활동 범위에 따라 동의하는 정도를 5점 리커트 척도로 측정하였다.

(3) 선호하는 개인정보 활용에 대한 동의 방식

사업자가 빅데이터 분석을 위해 개인정보를 활용할 때 소비자가 선택권을 발휘할 수 있는 방식을 제시하고 선호하는 것을 선택하도록 하였다. 사업자의 개인정보 활용에 동의하는 시점과 횟수 두 가지로 나누어 측정하였다.

먼저, ‘온라인 사이트에 가입하는 당시’와 ‘가입 후’ 등 두 시점으로 나누어 개인정보 활용에 대한 선택을 할 때 어느 시점을 선호하는지 택하도록 하였다. 또한 개인정보 활용에 대한 선택권을 발휘하는 횟수를 알아보기 위해 ‘단 한 번의 선택’과 ‘사업자가 개인정보를 활용할 때마다’의 두 가지 선택지를 제시하였다.

3. 조사도구의 신뢰도

Cronbach's α 값을 이용하여 측정 도구들의 신뢰도를 분석한 결과는 [표 3-3]과 같다.

먼저, 소비자들의 일반적인 개인정보와 관련된 변수들의 신뢰도 분석 결과를 살펴보면, 개인정보 보호행동을 측정하기 위한 4개 문항의 Cronbach's α 값이 0.69로 나타났다.

빅데이터 환경에서의 개인정보와 관련된 변수들의 신뢰도를 분석한 결과는 다음과 같다. 빅데이터 환경에서 생산되는 정보를 자신의 개인정보로 인식하는 정도를 측정한 문항들에 대한 신뢰도를 살펴보면, 자신이 전체 공개한 정보에 대한 인식의 Cronbach's α 값이 0.84, 타인이 전체 공개한 정보에 대한 인식은 0.87, 이용내역정보에 대한 인식은 0.91로 나타났다. 빅데이터 환경에서 개인정보가 활용될 때 소비자의 의사가 어느 정도 반영될 것인지에 대한 인식을 측정한 2개 문항의 Cronbach's α 값은 0.93으로 가장 높게 나타났다. 개인정보 활용에 대한 동의정도 문항들의 신뢰도를 분석한 결과, 공개된 정보의 활용에 대한 동의정도는

Cronbach's α 값이 0.83, 이용내역정보는 0.81이었다. 또한 사업자들이 빅데이터 환경에서의 개인정보를 활용하여 행하는 활동 범위에 따른 동의 정도는 0.85로 나타났다.

[표 3-3] 조사도구의 신뢰도

변수		Cronbach's α
일반적 개인정보 관련 변수		
개인정보 보호행동	정보공개 거부행동	0.69
	정책 확인행동	
빅데이터 환경에서의 개인정보 관련 변수		
빅데이터 환경에서의 개인정보에 대한 인식	자신이 공개한 정보	0.84
	타인이 공개한 정보	0.87
	이용내역정보	0.91
빅데이터 환경에서의 개인정보 활용에 대한 의사반영		0.93
빅데이터 환경에서의 개인정보 활용에 대한 동의정도	공개된 정보	0.83
	이용내역정보	0.81
	소극적 절차	0.87
	적극적 절차	0.89
	활동 범위	0.85

4. 분석방법

본 연구는 SPSS 프로그램을 사용하여 예비조사와 본조사를 통해 수집한 총 574명의 자료를 분석하였다.

먼저 조사대상자의 일반적 특성과 빅데이터 환경에서의 개인정보 및 개인정보의 활용에 대한 전반적 인식수준을 살펴보기 위해 빈도분석과 기술통계를 이용하였다. 또한 개인정보와 관련된 소비자의 행동, 경험에 따라 인식수준이 어떻게 달라지는지 알아보기 위해 t-test와 일원배치분산분석(ANOVA)을 실시하였다. 분산분석을 실시한 이후에 사후검정(Duncan)을 실시하여 집단별 차이가 통계적으로 유의한지 살펴보았다.

사업자가 빅데이터 환경에서 생산된 개인정보를 활용하는데 소비자가 동의하는 정도에 영향을 미치는 요인을 알아보기 위해 다중회귀분석을 실시하였다. 사업자가 개인정보를 활용할 때 취하는 조치에 따른 동의정도와 개인정보 범주별 동의정도에 영향을 미치는 요인을 분석하였다.

제 3 절 조사대상자의 일반적 특성

본 연구에서 조사를 통해 결과 분석에 쓰인 표본은 총 574명이다. 조사대상자의 일반적 특성은 [표 3-4]와 같다. 먼저 성별을 살펴보면, 남자가 282명(49.1%), 여자가 292명(50.9%)로 여자가 남자보다 조금 더 많았다. 연령은 만 14-19세인 10대가 111명(19.3%), 만 20-29세인 20대가 121명(21.1%), 30대가 111명(19.3%), 40대가 119명(20.7%), 50대가 112명(19.5%)이었다. 조사대상자의 직업은 사무·기술직이 226명(39.4%)로 가장 많았고, 학생이 163명(28.4%), 주부가 63명(11.0%)이었으며, 그 다음으로 자영업 33명(5.7%), 판매·서비스직 27명(4.7%), 기타 24명(4.2%), 생산·기능·노무직 20명(3.5%), 기업인·경영직 18명(3.1%) 순이었다. 최종학력은 대졸/전문대졸(중퇴 및 재학 포함)이 376명(65.5%)로 가장 많았고, 고졸(중퇴 포함)이 129명(22.5%), 대학원 이상(재학 포함) 55명(9.6%), 중졸 이하 14명(2.4%)로를 차지했다. 조사대상자가 속한 가계의 월평균 소득은 300만원 이상부터 500만원 미만까지 215명(37.5%)으로 가장 많았고, 100만원 이상부터 300만원 미만까지 157명(27.4%), 500만원 이상에서 700만원 미만까지 104명(18.1%), 700만원 이상은 74명(12.9%), 100만원 미만은 24명(4.2%)이었다.

[표 3-4] 조사대상자의 일반적 특성 (n=574)

특성	구분	빈도(명)	백분율(%)
성별	남자	282	49.1
	여자	292	50.9
연령	10대(만 14-19세)	111	19.3
	20대(만 20-29세)	121	21.1
	30대(만 30-39세)	111	19.3
	40대(만 40-49세)	119	20.7
	50대(만 50-59세)	112	19.5
직업	기업인·경영직	18	3.1
	사무·기술직	226	39.4
	판매·서비스직	27	4.7
	자영업	33	5.7
	생산·기능·노무직	20	3.5
	학생	163	28.4
	주부	63	11.0
	기타	24	4.2
최종학력	중졸 이하	14	2.4
	고졸(중퇴 포함)	129	22.5
	대졸/전문대졸(중퇴 및 재학 포함)	376	65.5
	대학원 이상(재학 포함)	55	9.6
월평균소득	100만원 미만	24	4.2
	100만원 이상~300만원 미만	157	27.4
	300만원 이상~500만원 미만	215	37.5
	500만원 이상~700만원 미만	104	18.1
	700만원 이상	74	12.9
계		574	100.0

제 4 장 연구결과

본 장은 앞장에서 제시한 연구문제에 대한 분석 결과이다. 제1절에서는 빅데이터 환경에서의 개인정보에 대한 인식을 살펴보고, 제2절에서는 빅데이터 환경에서의 개인정보 활용에 대한 인식을 살펴보았다.

제 1 절 빅데이터 환경에서의 개인정보에 대한 인식

본 절에서는 소비자가 빅데이터 환경에서 생산되는 정보를 얼마나 자신의 개인정보라고 인식하는지에 대해 알아보고, 개인정보와 관련된 행동, 경험 및 기타 인구통계학적 변수에 따라 인식 수준이 어떠한 차이를 보이는지 알아보았다.

1. 빅데이터 환경에서의 개인정보에 대한 인식수준

본 연구에서 빅데이터 환경에서의 개인정보에 대한 인식은 ‘빅데이터 환경에서 발생하는 정보를 자신의 개인정보라고 생각하는 정도’를 의미한다. 기존과는 달리 온라인 사이트 가입 후 소비자들이 적극적으로 생산하는 정보가 사업자에 의해 수집되고 활용되어 가치를 생산한다. 정보들에 대한 소비자들의 근본적인 인식, 즉 빅데이터 환경에서 발생하는 정보를 자신의 개인정보라고 생각하는지 살펴보았다.

먼저 정보를 개인정보보호 가이드라인에서 발표한 공개된 정보와 이용내역정보로 분류하였다. 공개된 정보는 공개주체별로 자신과 타인으로 구분하여 그에 속하는 개인정보들을 나열하였는데, 구체적으로 공개된 정보에 속한 개별 정보들은 페이스북 등 소셜네트워크서비스에서 소비자가 전체 공개할 수 있는 정보들로 구성하였다.

구체적으로 ‘나는 (_____)을 나의 개인정보라고 생각한다’라는 질문

으로 공개된 정보 10개, 이용내역정보 5개, 총 15개의 개인정보에 대해 5점 리커트 척도(전혀 그렇지 않다-별로 그렇지 않다-보통이다-조금 그렇다-매우 그렇다)로 측정하였다. 개별 정보에 대해 소비자가 자신의 개인정보라고 인식하는 정도를 알아본 결과는 [표 4-1]과 같다. 정보별 정확한 순위를 살펴보기 위해 5개의 정보에 대한 평균은 소수점 세 자리까지 표시하였다.

[표 4-1] 빅데이터 환경에서의 정보를 개인정보로 인식하는 수준

순위	빅데이터 환경에서의 정보 종류	평균(SD)
1	자신이 공개한 일반 인적 사항	4.284(.93)
2	자신이 공개한 위치정보	4.281(.92)
3	사이트 접속 실시간 기록 (현재 위치 정보 등)	4.275(.87)
4	사이트 거래기록 (쇼핑 내역 등)	4.24(.87)
5	타인이 공개한 내가 포함된 위치정보	4.17(.87)
6	자신이 공개한 사진 및 동영상	4.16(.95)
7	타인이 공개한 내가 포함된 사진 및 동영상	4.14(.85)
8	자신이 공개한 온라인 인적 사항	4.09(.92)
9	사이트 접속기록 (접속 시간, 장소 등)	4.06(.90)
10	타인이 공개한 내가 포함된 사회적 정보	4.038(.84)
11	자신이 공개한 사회적 정보	4.035(.94)
12	사이트 검색기록 (검색어 등)	3.98(.96)
13	사이트 클릭기록 (광고, 방문 사이트 등)	3.91(.96)
14	타인이 공개한 내가 포함된 글	3.68(.90)
15	자신이 공개한 글	3.54(.96)

모든 정보들에 대해 소비자들이 자신의 개인정보라고 인식하는 수준이 3점을 넘었으며, 11개의 정보에 대해서는 4점 이상으로 매우 높은 수준임을 알 수 있다. 소비자가 자신의 개인정보라고 인식하는 정도가 높은 정보의 순서별로 살펴보면 다음과 같다. 먼저 이름, 성별, 학력, 직업 등을 포함하는 자신이 전체 공개한 일반 인적 사항을 개인정보라고 가장 높게 인식하고 있다. 다음으로 개인정보로 인식되는 정보 2위는 자신이 공개한 위치정보였다. 실시간으로 파악되는 현재 위치 정보 등과 같이 사이트 접속 시 수집되는 실시간 기록이 3위, 쇼핑 내역 등과 같은 사이트 거래기록이 4위로, 소비자들이 이러한 정보들을 자신의 개인정보라고 높게 인식하였다. 타인이 전체 공개한 자신이 포함된 위치정보가 5위를 차지하였다.

공개된 정보에 속하는 자신과 타인이 공개한 자신이 포함된 사진 및 동영상은 각각 4.16과 4.14점으로 개인정보로 인식되는 순위 6, 7위를 차지하였다. 온라인 상에서 특정한 개인을 대표하는 아이디나 이메일 등 온라인 인적 사항은 8위였다. 사이트에 접속한 시간과 장소 등 접속기록이 4.06점으로 9위에 올랐다. SNS 인맥 및 동호회 가입 등을 드러내는 온라인 상에 전체 공개된 사회적 정보는 근소한 차이로 정보 공개 주체에 따라 타인이 공개한 정보는 10위, 자신이 공개한 정보는 11위였다.

온라인 사이트에서 자신이 입력한 검색어, 광고를 포함해 자신이 클릭하여 방문한 사이트에 대한 기록 등 사이트 검색 및 클릭기록은 12위, 13위를 차지하였다. 타 정보들에 비해 타인 및 자신이 전체 공개한 글은 개인정보라고 인식되는 정도가 낮게 나타났다. 사진 및 동영상은 유포되었을 경우 이미지로 자신을 한 번에 알아볼 수 있는 특징을 지니고 있어 글보다 더 민감할 수 있는데, 이와 같은 특성으로 인해 인식의 차이가 발생한 것으로 유추할 수 있다.

전반적으로 소비자들이 빅데이터 환경에서 생성되는 정보들을 자신의 개인정보라고 높게 인식하고 있었다. 정보 종류별로 인식의 차이도 나타났는데 특히 위치정보를, 또한 글보다는 사진 및 동영상을 자신의 개인정보로 생각하는 정도가 높음을 알 수 있다.

정보 범주별 개인정보로 인식하는 정도의 차이는 [표 4-2]와 같다. 공개된 정보를 주체별로 구분한 자신이 전체 공개한 정보와 타인이 전체 공개한 정보에 대한 인식의 차이는 통계적으로 유의하지 않았다. 하지만 전체 공개한 정보와 이용내역정보 간 인식의 차이는 유의한 차이를 보였다. 즉, 자신 및 타인이 전체 공개한 정보보다는 이용내역정보를 자신의 개인정보라고 인식하고 평균 점수가 더 높았다.

[표 4-2] 정보 범주별 개인정보로 인식하는 정도의 차이

정보 범주	평균(표준편차)	t	
자신이 전체 공개한 정보	4.00(.75)	3.18**	
이용내역정보	4.09(.78)		
타인이 전체 공개한 정보	4.01(.74)		-2.93**

유의수준: ** p<.01

2. 빅데이터 환경에서의 개인정보에 대한 인식 차이

1) 인구통계학적 특성에 따른 인식 차이

빅데이터 환경에서의 정보를 위에서 살펴본 세 가지 범주로 구분하여 인구통계학적 특성에 따라 정보에 대한 인식의 평균 차이가 유의한지 살펴보았다. 빅데이터 환경에서의 정보를 개인정보로 인식하는 정도에 차이가 있는지 알아본 결과는 [표 4-3]와 같다.

인구통계학적 특성 중 성별과 연령대, 직업에 따라 빅데이터 환경에서 생성되는 개인정보 범주별(공개된 정보, 이용내역정보) 평균의 차이가 나타났다. 남성보다는 여성이 이용내역정보를 자신의 개인정보라고 더 높게 인식하였다. 또한 다른 연령대에 비해 50대가 빅데이터 환경에서 발생하는 정보를 자신의 개인정보라고 생각하는 정도가 낮았다. 하지만 다른 인구통계학적 특성인 최종학력과 월평균 소득에 따른 개인정보에 대한 인식의 차이는 보이지 않았다.

상대적으로 남성에 비해 여성이, 또한 젊은 연령층이 SNS와 인터넷 쇼핑 이용률이 높는데(한국인터넷진흥원, 2014), 이용률의 차이가 빅데이터 환경에서 생성되는 개인정보에 대한 인식에 영향을 미친 것으로 유추할 수 있다. 즉, SNS에 자신에 관한 정보를 자주 게재하고 쇼핑몰에 자신의 흔적이 많이 기록되는 소비자가 빅데이터 환경에서의 개인정보에 관심을 가지고 민감해질 가능성이 존재한다.

[표 4-3] 인구통계학적 특성에 따른 인식 차이 (n=574)

구분		빈도	공개된 정보(자신)			공개된 정보(타인)			이용내역정보		
			평균(SD)	D	t/F	평균(SD)	D	t/F	평균(SD)	D	t/F
성별	남	282	3.88(.78)		-3.82	3.90(.75)		-3.60	4.01(.81)		-2.37*
	여	292	4.12(.71)			4.12(.71)			4.17(.74)		
연령	10대	111	4.03(.78)	b	5.94***	4.05(.75)		2.03	4.20(.74)	b	3.76**
	20대	121	4.11(.75)	b		4.03(.72)			4.16(.85)	b	
	30대	111	4.06(.72)	b		4.00(.77)			4.17(.76)	b	
	40대	119	4.10(.63)	b		4.10(.70)			4.08(.74)	b	
	50대	112	3.70(.82)	a		3.84(.75)			3.85(.75)	a	
최종 학력	고졸이하	143	3.92(.84)		1.01	4.08(.76)		1.02	4.10(.79)		.01
	대졸	376	4.03(.72)			3.98(.73)			4.09(.79)		
	대학원이상	55	4.02(.75)			3.97(.73)			4.10(.69)		
직업	학생	163	4.05(.77)		.88	4.04(.74)		.66	4.22(.75)		2.45*
	학생 외	411	3.98(.75)			4.00(.74)			4.04(.78)		
월평균 소득	300만원 미만	181	3.96(.79)		1.12	3.98(.77)		1.14	4.05(.79)		.33
	300만원 이상- 500만원 미만	215	3.98(.74)			3.98(.74)			4.10(.80)		
	500만원 이상	178	4.07(.73)			4.08(.70)			4.12(.74)		

유의수준: *** p<.001, ** p<.01, * p<.05

2) 개인정보 관련 소비자의 행동, 경험에 따른 인식 차이

개인정보 관련 소비자의 행동, 경험에 따라 빅데이터 환경에서 생산되는 정보를 개인정보로 인식하는 정도에 차이가 있는지 정보 범주별로 구분하여 살펴보았다. 개인정보 관련 행동에는 일반적으로 개인정보를 보호하기 위해 취하는 행동과 온라인 사이트에 가입하는 당시 약관을 확인하는 행동 두 가지가 있다.

소비자들의 전반적인 개인정보 관련 행동과 경험의 현황은 다음과 같다. 먼저 소비자가 평소에 개인정보를 보호하기 위해 어느 정도의 행동을 취하고 있는지 알아본 결과 평균 수준은 3.43점으로 보통보다 약간 높았다. 그 중 정보공개 거부 행동의 평균 수준이 3.81점으로 개인정보관련 정책 및 사안을 확인하는 행동의 평균 수준인 3.05점보다 높았다.

소비자가 온라인 사이트에 가입할 시 개인정보 수집·활용 약관을 어느 정도 확인하고 있는지 살펴본 결과, 내용은 읽어보지 않고 약관 동의 시 체크해야 하는 부분만 찾아서 표시하는 소비자가 234명으로 40.8%에 달했고, 대충 스크롤을 내려보고 체크하는 소비자가 213명으로 37.1%를 차지했다. 즉, 개인정보 수집·활용 약관에 동의 시 내용을 읽어보지 않는 소비자가 77.9%로 대부분이었다. 이에 반해 약관을 조금이라도 읽어보는 소비자가 104명으로 18.1%였고, 꼼꼼히 읽어보는 사람은 23명으로 4%를 차지하였다. 즉, 약관의 내용을 읽어보는 소비자는 조사대상자의 22.1% 밖에 되지 않았다.

다음으로 소비자가 개인정보 유출을 경험한 적이 있는지 알아보았다. 그 결과 유출 경험이 명확하게 '없다'라고 답한 소비자가 80명으로 13.9%였고, 유출 경험이 '있다'라고 답한 소비자는 절반이 넘는 294명으로 51.2%였다. 반면 개인정보의 유출을 경험했는지 알 수 없다고 답한 소비자는 200명으로 34.8%를 차지하였다.

개인정보 관련 행동과 경험에 따른 개인정보에 대한 인식에 차이가 있는지 살펴보기 위해 개인정보 보호행동 점수가 3점 미만인 집단을 보호행동이 낮은 집단, 3점인 집단을 중간 집단, 3점 초과인 집단을 높은 집

단으로 구분하였다. 약관 확인행동은 약관을 읽는 집단과 읽지 않는 집단으로 구분하여 이에 따른 개인정보에 대한 인식의 차이를 살펴본 결과는 [표 4-4]와 같다.

개인정보 보호행동 점수가 높은 집단이 다른 집단에 비해 자신 및 타인이 전체 공개한 정보와 이용내역정보를 자신의 개인정보라고 인식하는 수준이 높았다. 하지만 약관을 확인하는 행동에 따른 개인정보 인식의 차이는 통계적으로 유의하지 않았다. 또한 개인정보 유출 경험이 없는 사람보다 유출 경험이 있거나 자신의 개인정보가 유출되었는지 잘 모르는 소비자가 빅데이터 환경에서 생성된 정보를 자신의 개인정보라고 높게 생각하였다.

개인정보 보호행동 점수가 높은 집단이 일반적으로 개인정보에 더 민감하여 빅데이터 환경에서 생산되는 정보 또한 자신의 개인정보로 높게 인식함을 유추할 수 있다. 또한 개인정보 유출 경험이 있거나 이에 대해 확신하지 못하는 집단이 개인정보에 대한 우려가 크고 민감하다고 볼 수 있는데, 이 특성이 개인정보에 대한 인식에 영향을 미친 것으로 유추할 수 있다.

[표 4-4] 개인정보 관련 행동과 경험에 따른 인식 차이 (n=574)

구분		빈도	공개된 정보(자신)			공개된 정보(타인)			이용내역정보
			평균(SD)	D	t/F	평균(SD)	D	t/F	
개인정보 보호행동	낮음	190	3.83(.79)	a	13.07***	3.82(.74)	a	14.92***	3.97(.88)
	중간	230	3.99(.70)	b		4.00(.70)	b		4.08(.70)
	높음	154	4.24(.72)	c		4.25(.72)	c		4.26(.72)
약관 확인행동	읽음	127	4.20(.76)	1.70	4.06(.73)	.92	4.15(.77)		
	읽지 않음	447	3.97(.75)		3.99(.74)		4.08(.78)		
개인정보 유출경험	없다	80	3.83(.71)	3.02	3.77(.73)	a	5.45**	3.85(.82)	
	있다	294	4.06(.75)		4.08(.74)	b		4.14(.77)	
	모르겠다	200	3.99(.77)		4.00(.72)	b		4.11(.76)	

유의수준: *** p<.001, ** p<.01

제 2 절 빅데이터 환경에서의 개인정보 활용에 대한 인식

본 절에서는 사업자가 빅데이터 환경에서 생성되는 개인정보를 활용하는 것에 대해 소비자가 어떻게 인식하고 있는지 살펴보았다. 나아가 소비자가 사업자의 개인정보 활용에 동의하는 정도에 영향을 미치는 변수는 무엇인지 알아보고자 한다.

1. 빅데이터 환경에서의 개인정보 활용에 대한 인식수준

1) 개인정보 활용에서의 의사 반영에 대한 생각

빅데이터 환경에서 생성되는 개인정보를 사업자가 활용하는데 있어서 소비자의 의사가 어느 정도 반영될 것이라고 생각하는지를 측정하기 위해 “빅데이터 환경에서의 개인정보가 활용되는 과정/범위에 자신의 의사가 충분히 반영될 것이다”라는 질문을 하였다. 이에 대한 소비자의 응답은 “전혀 그렇지 않다-별로 그렇지 않다-보통이다-조금 그렇다-매우 그렇다” 다섯 가지로, 리커트 5점 척도를 이용하여 측정하였다.

사업자가 빅데이터 분석을 위해 개인정보를 활용하는 과정 및 범위에 소비자의 의사가 어느 정도 반영될 지에 관한 인식의 분포를 보여주는 결과는 다음과 같다. 먼저 개인정보가 활용되는 과정에서 자신의 의사가 ‘전혀 반영되지 않을 것이다’라고 생각하는 소비자가 51명(8.9%)이었고, ‘별로 반영되지 않을 것이다’라고 생각하는 소비자가 191명(33.3%), ‘보통이다’라고 응답한 소비자가 190명(33.1%)로 나타났다. 이에 반해 의사가 ‘조금 반영될 것이다’라고 답한 소비자는 114명(19.9%), ‘매우 그렇다’가 28명(4.8%)이었다.

마찬가지로 개인정보가 활용되는 범위에 자신의 의사가 반영될지에 관한 인식의 분포도 비슷한 양상을 보였다. 개인정보가 활용되는 범위에서 자신의 의사가 ‘전혀 반영되지 않을 것이다’라고 생각하는 소비자가 57명(9.9%)이었고, ‘별로 반영되지 않을 것이다’라고 생각하는 소비자가 176

명(30.7%), ‘보통이다’라고 응답한 소비자가 190명(33.1%)로 나타났다. 이에 반해 의사가 ‘조금 반영될 것이다’라고 답한 소비자는 119명(20.7%), ‘매우 그렇다’가 32명(5.6%)이었다.

빅데이터 환경에서 개인정보가 활용되는 과정에 소비자의 의사가 반영될 것이라고 생각하는 정도는 5점 만점의 리커트 척도에서 2.79점, 정보가 활용되는 범위에서 의사가 반영될 것이라고 생각하는 정도는 2.81점으로 나타났다. 즉 빅데이터 분석을 위해 개인정보가 처리될 때 많은 소비자가 자신의 의견이 반영되지 않을 것이라고 답하였다. 도출된 결과는 빅데이터 환경에서 사업자가 개인정보를 활용할 때 소비자 자신이 발휘할 수 있는 통제력이 강하지 않을 것이라 인식하고 있음을 반영한다.

2) 빅데이터 환경에서의 개인정보 활용에 대한 소비자의 동의정도

(1) 개인정보 범주별 활용에 대한 동의정도

사업자가 빅데이터 환경에서 생산된 개인정보를 활용할 때 소비자가 어느 정도 동의하는지 알아보았다. 설문에 앞서 빅데이터 개인정보보호 가이드라인에서 개인정보의 보호를 위해 강조한 비식별화 조치에 대한 설명을 추가하였다. ‘사업자가 나의 개인정보를 모두 내가 누군지 알 수 없게 조치를 한 상태입니다’라는 문구를 통해 비식별화 조치가 이루어진 후에 개인정보가 활용된다는 사실을 소비자들이 인지할 수 있도록 설문 문항을 구성하였다.

사업자가 빅데이터 환경에서 발생한 개인정보를 비식별화한 후 이를 활용하는데 있어 소비자의 동의정도가 어떠한지 살펴본 결과는 [표 4-5]와 같다. 개인정보 범주별 응답 분포를 살펴본 결과 공개된 정보와 이용내역정보가 비슷한 양상을 보였다. 먼저 공개된 정보를 활용할 때 사업자 별다른 조치를 취하지 않는 경우 개인정보 활용에 동의하지 않는 소비자, 즉 ‘전혀 동의하지 않는다’와 ‘별로 동의하지 않는다’라고 답한 사람들이 총 358명으로 62.4%를 차지하였다. 사업자가 공개된 정보를 활용한다는 사실을 쉽게 확인할 수 있도록 공개만 했을 경우에도 304명인

55.9%의 소비자들이 개인정보의 활용에 동의하지 않는 것으로 나타났다. 이와 달리 사업자가 개인정보 활용에 대해 소비자에게 고지하거나 동의를 받을 경우에는 ‘보통이다’, ‘조금 동의한다’라고 답한 비율이 증가하였음을 알 수 있다.

이용내역정보의 활용에 대한 동의정도의 분포를 살펴보면 소비자가 온라인 사이트에 가입하는 것만으로, 즉 사업자의 특별한 조치가 없는 경우 혹은 사업자가 개인정보의 활용에 대해 소비자가 쉽게 확인할 수 있도록 조치를 취했을 경우에는 ‘전혀 동의하지 않는다’와 ‘조금 동의하지 않는다’라고 답한 비율이 60%에 달하였다. 이와 달리 사업자가 소비자에게 개인정보 활용에 대해 고지하거나 동의를 받는 조치를 취할 경우 동의하는 정도가 높아짐을 알 수 있다.

전반적으로 사업자가 빅데이터 환경에서 발생한 개인정보를 활용할 때 소비자에게 적극적으로 알리고 동의를 받는 경우 정보 활용에 동의하는 정도가 높아짐을 알 수 있다. 이는 사업자가 취하는 조치별 동의하는 정도의 평균을 통해서도 확인할 수 있다. 정보의 종류와 무관하게 사업자가 개인정보를 활용할 때 소극적 조치를 취하면 동의정도가 2.23점으로 낮게 나타났다. 하지만 적극적인 조치를 취하였을 경우 동의정도가 3.15점으로 상대적으로 높아짐을 알 수 있다($t=-24.64, p<.001$).

정보별 동의 정도를 살펴보면 공개된 정보에 대한 동의정도는 5점 만점에서 2.80점으로 나타났고, 이용내역정보에 대한 동의정도는 2.59점이었다. 즉, 소비자들은 공개된 정보보다 통제하기 어려운 이용내역정보의 활용에 더 동의하지 않았다($t=8.05, p<.001$). 공개된 정보의 경우 소비자가 생산하고 공개하는 정보의 양을 조절할 수 있다. 하지만 온라인 사이트를 이용하는 과정에서 자동적으로 발생하는 이용내역정보의 경우에는 그 정보가 생산되고 수집되는 과정을 소비자가 통제하기 어렵기 때문에 이러한 결과가 나타난 것으로 유추할 수 있다.

[표 4-5] 개인정보 범주별 활용에 대한 동의정도 (n=547)

빈도(%)		전혀 그렇지 않다	별로 그렇지 않다	보통 이다	조금 그렇다	매우 그렇다	평균 (SD)
공 개 된 정 보	별다른 조치 없어도	153	205	146	59	11	2.25
	정보 활용에 동의	(26.7)	(35.7)	(25.4)	(10.3)	(1.9)	(1.02)
	쉽게 확인할 수 있으면	131	173	168	87	15	2.45
	정보 활용에 동의	(22.8)	(30.1)	(29.3)	(15.2)	(2.6)	(1.08)
	자신에게 알리면	70	116	174	171	43	3.00
	정보 활용에 동의	(12.2)	(20.2)	(30.3)	(29.8)	(7.5)	(1.14)
이 용 내 역 정 보	자신에게 동의를 받는다면	42	67	145	206	114	3.49
	정보 활용에 동의	(7.3)	(11.7)	(25.3)	(35.9)	(19.9)	(1.15)
	가입하는 것만으로도	226	184	120	36	8	1.98
	정보 활용에 동의	(39.4)	(32.1)	(20.9)	(6.3)	(1.4)	(.99)
	쉽게 확인할 수 있으면	169	174	157	65	9	2.25
	정보 활용에 동의	(29.4)	(30.3)	(27.4)	(11.3)	(1.6)	(1.05)
정 보	자신에게 알리면	91	139	174	136	34	2.80
	정보 활용에 동의	(15.9)	(24.2)	(30.3)	(23.7)	(5.9)	(.14)
	자신에게 동의를 받는다면	62	71	158	189	94	3.32
정보 활용에 동의	(10.8)	(12.4)	(27.5)	(32.9)	(16.4)	(1.20)	

(2) 사업자의 개인정보 활용 범위에 따른 동의정도

사업자가 빅데이터 환경에서 생산되는 개인정보를 활용하여 어떠한 활동을 하는지에 따라 소비자의 동의정도에 차이가 발생하는지 살펴보았다. 빅데이터 개인정보를 활용하여 메일과 문자, 전화 등을 통해 소비자에게 광고 활동을 하는 것에 동의하는지 알아본 결과 ‘전혀 그렇지 않다’와 ‘별로 그렇지 않다’라고 답한 소비자들이 61.1%를 차지했다.

사업자가 소비자에게 ‘맞춤형 광고’ 혹은 ‘제품 개발’과 같은 장·단기적인 혜택을 제공할 경우에 빅데이터 개인정보의 활용에 어느 정도 동의하는지 살펴보았다. 맞춤형 광고를 제공할 때 동의하는 정도가 ‘보통이다’라고 답한 소비자들이 36.1%였고, ‘조금 그렇다’와 ‘매우 그렇다’라고 답한 소비자들이 21.1%를 차지하였다. 제품 개발 및 전반적인 홍보 활동에 동의하는 정도가 ‘보통이다’인 소비자가 36.8%였으며, ‘조금 그렇다’와 ‘매우 그렇다’라고 답한 소비자들이 17.7%였다.

소비자들의 동의정도의 평균을 살펴본 결과 전반적으로 3점보다 낮았으며, 이는 소비자들이 사업자가 빅데이터 환경에서 발생한 개인정보를 활용하는 것에 동의하지 않음을 보여준다. 그러나 사업자가 빅데이터 개인정보를 활용하여 소비자에게 장·단기적 혜택을 제공할 때에는 동의하는 정도가 상대적으로 높아짐을 알 수 있다.

[표 4-6] 사업자의 개인정보 활용 범위에 따른 동의정도 (n=547)

빈도(%)	전혀 그렇지 않다	별로 그렇지 않다	보통 이다	조금 그렇다	매우 그렇다	평균 (SD)
메일, 문자, 전화 등 광고	169 (29.4)	182 (31.7)	146 (25.4)	68 (11.8)	9 (1.6)	2.24 (1.05)
맞춤형 광고 제공	99 (17.2)	147 (25.6)	207 (36.1)	104 (18.1)	17 (3.0)	2.64 (1.06)
제품 개발 및 전반적 홍보	110 (19.2)	151 (26.3)	211 (36.8)	88 (15.3)	14 (2.4)	2.56 (1.04)

3) 개인정보 활용에 대한 동의가 필수일 경우 사이트 가입의사

온라인 사이트의 사업자가 가입 단계부터 빅데이터 환경에서 생산되는 개인정보를 활용하기 위해 동의를 받는다는 전제를 바탕으로 질문을 구성하였다. 이용하고자 하는 사이트에 가입 시 빅데이터 분석을 위한 개인정보 수집·활용에 동의하지 않으면 가입할 수 없을 경우 소비자들의 가입 의사를 알아보았다.

사업자의 개인정보 활용에 무조건 동의하고 가입하겠다는 소비자가 57명으로 9.9%였다. 하지만 ‘가입을 고려하겠다’ 즉, ‘가입을 유보하겠다’는 소비자들이 457명으로 79.6%를 차지하였고, 무조건 가입하지 않겠다고 답한 사람들이 60명으로 10.5%를 차지하였다. 이를 통해 응답자들의 90% 이상이 해당 사이트에 가입을 하지 않거나 유보할 만큼 전반적으로 소비자들이 온라인 사이트에서 사업자가 빅데이터 분석을 위해 개인정보를 활용하는 것을 원하지 않고 있음을 유추할 수 있다.

4) 선호하는 개인정보 활용에 대한 동의 방식

사업자가 빅데이터 환경에서의 개인정보를 활용하는 것에 소비자가 선택권을 지닐 수 있을 때 어떠한 절차를 통해 활용에 동의하는 것을 선호하는지 살펴본 결과는 [표 4-7]과 같다.

먼저 온라인 사이트에 가입하는 당시에 사업자의 개인정보 활용에 대해 선택하기를 선호하는 소비자가 71.1%를 차지하였다. 즉, 과반수가 넘는 소비자들이 사이트 가입 후보다는 가입 시에 개인정보 활용에 대한 선택권을 발휘하길 원하였다.

빅데이터 환경에서 발생하는 정보의 특성상 사업자가 개인정보를 빈번하게 수집하고 활용할 수 있다. 사업자가 개인정보를 여러 번 활용한다는 것을 소비자가 인지하고 있을 때 선호하는 활용에 대한 선택의 횟수를 알아본 결과는 다음과 같다. 사업자가 개인정보를 활용할 때마다 매번 선택할 수 있기를 원하는 소비자들이 과반수를 넘는 68.3%를 차지하였다. 이와 달리 개인정보에 활용에 대한 선택을 단 한번으로 원하는 소

비자는 31.7%였다.

소비자가 선택한 개인정보 활용에 대한 동의 방식을 살펴본 결과, 전반적으로 사업자의 개인정보 활용에 대해 소비자들이 적극적인 선택권을 발휘하기를 선호하였다. 빅데이터 환경에서 다양한 개인정보가 실시간으로 생산되어 정보의 양이 방대하기 때문에 현실적으로 모든 정보에 대해 소비자의 동의를 받는 것은 어렵다. 하지만 빅데이터 개인정보보호 가이드라인에서 제시하고 있는 ‘정보 활용 거부권 행사 페이지’는 소비자의 선호를 반영하지 못하는 소극적인 대안이므로 이를 수정할 필요가 있다.

[표 4-7] 선호하는 개인정보 활용에 대한 동의 방식

방식		빈도(%)
가입 시점	사이트 가입 시 개인정보 활용에 대한 선택	408(71.1)
	사이트 가입 후 개인정보 활용에 대한 선택	166(28.9)
	계	574(100)
횟수	단 한 번으로 개인정보 활용에 대한 선택	182(31.7)
	개인정보를 활용할 때마다 매번 활용에 대한 선택	392(68.3)
	계	574(100)

2. 빅데이터 환경에서의 개인정보 활용에 대한 동의정도에 영향을 미치는 요인

1) 관련변수 간 다중공선성 문제 진단

사업자가 빅데이터 환경에서 생산되는 개인정보를 활용하는데 있어 소비자의 동의정도에 영향을 미치는 요인을 살펴보기 위해 회귀분석을 실시하였다. 회귀분석을 실시하기 전 다중공선성 문제를 진단하고자 독립변수 간 상관관계와 분산팽창요인(VIF)를 살펴보았다.

그 결과 독립변수 상호 간에 유의한 상관관계가 발견되었으나 모든 상관관계수의 절대값이 0.3 이하로 낮았다. 공차한계와 분산팽창요인(VIF)을 살펴본 결과 공차한계는 모두 0.1보다 높았으며, 분산팽창요인의 최대값은 2.45로 다중공선성의 문제가 없음을 확인하였다.

2) 개인정보 활용에 대한 소비자의 동의정도 다중회귀분석 결과

사업자가 빅데이터 환경에서 생산되는 개인정보를 활용하는데 동의하는 정도에 영향을 미치는 요인을 알아보기 위해 다중회귀분석을 실시한 결과는 [표 4-8]과 같다.

독립변수를 4가지로 분류하였다. 먼저 개인정보별 활용 동의는 공개된 정보와 이용내역정보의 활용에 대한 동의로 구분하였다. 또한 사업자의 조치별 활용 동의는 사업자가 소극적 절차를 취할 때와 적극적 절차를 통해 정보에 대한 동의를 받을 경우로 살펴보았다. 개인정보별 활용 동의 2개와 사업자 조치별 활용 동의 2개, 총 4개의 회귀식은 모두 통계적으로 유의하였다.

‘공개된 정보의 활용에 대한 동의정도’를 종속변수로 하는 회귀식 ($F=4.50, p<.001$)에서 유의한 영향을 미치는 요인으로는 직업($b=.37, p<.01$), 개인정보 보호행동($b=-.20, p<.001$), 의사반영에 대한 인식($b=.15,$

$p < .001$)이 있다. 이는 학생이 아닌 사람보다는 학생이 공개된 정보의 활용에 더 동의함을 의미한다. 사업자가 빅데이터 환경에서 생산되는 개인 정보를 활용할 때 소비자의 의사가 반영이 된다고 인식할수록 공개된 정보의 활용에 더 동의하는 것으로 나타났다. 이와 다르게 개인정보 보호 행동 점수가 높을수록 공개된 정보의 활용에 동의하지 않는 것으로 나타났다. 변수 간 영향력을 비교하였을 때 가장 큰 영향을 미치는 변수는 직업, 개인정보 보호행동과 활용 시 의사반영 순으로 나타났다.

다음으로 ‘이용내역정보의 활용에 대한 동의정도’를 종속변수로 하는 회귀식($F=5.35$, $p < .001$)에서 이용내역정보의 활용에 대한 동의정도에 영향을 미치는 요인에는 성별($b=.16$, $p < .05$), 연령($b=.01$, $p < .01$), 직업($b=.31$, $p < .05$), 개인정보 보호행동($b=-.12$, $p < .05$), 빅데이터 환경에서의 개인정보에 대한 인식($b=-.12$, $p < .05$), 의사반영에 대한 인식($b=.13$, $p < .001$)이 있다.

마찬가지로 ‘사업자의 소극적 절차 시 활용에 대한 동의정도’를 종속변수로 하는 회귀식($F=5.35$, $p < .001$)에 영향을 미치는 요인이 동일하였다. 즉, 사업자가 소극적 절차를 통해 활용에 대한 동의를 받을 때 소비자의 동의정도에 영향을 미치는 요인이 성별($b=.16$, $p < .05$), 연령($b=.01$, $p < .01$), 직업($b=.32$, $p < .01$), 개인정보 보호행동($b=-.11$, $p < .05$), 빅데이터 환경에서의 개인정보에 대한 인식($b=-.18$, $p < .01$), 의사반영에 대한 인식($b=.20$, $p < .001$)이었다.

결과를 통해 여성보다는 남성이, 연령이 높아질수록, 학생이 아닌 사람보다는 학생이 이용내역정보의 활용과 사업자의 소극적 절차를 통한 개인정보의 활용에 더 동의한다고 할 수 있다. 또한 사업자가 빅데이터 환경에서 생산된 개인정보를 활용할 때 소비자의 의사가 반영이 된다고 생각할수록 개인정보의 활용에 더 동의하는 것으로 나타났다. 이와 다르게 개인정보 보호행동 점수가 높을수록, 빅데이터 환경에서의 정보를 자신의 개인정보라고 생각할수록 정보의 활용에 동의하지 않았다.

하지만 두 회귀식에서 독립변수들 간 상대적 영향력의 차이가 존재했다. 먼저 ‘이용내역정보의 활용에 대한 동의정도’에 가장 큰 영향을 미치

는 변수는 연령과 직업, 의사반영에 대한 인식, 개인정보 보호행동, 성별과 빅데이터 환경에서의 개인정보에 대한 인식 순으로 나타났다. ‘사업자의 소극적 절차 시 활용에 대한 동의정도’에 가장 큰 영향을 미치는 변수는 의사반영에 대한 인식이었고, 다음으로 연령, 직업, 빅데이터 환경에서의 개인정보에 대한 인식 성별, 개인정보 보호행동 순으로 나타났다.

‘사업자의 적극적 절차 시 활용에 대한 동의정도’를 종속변수로 하는 회귀식($F=4.50, p<.001$)에서 사업자가 소극적 절차를 통해 활용에 대한 동의를 받을 때 소비자의 동의정도에 영향을 미치는 요인에는 직업 ($b=.36, p<.01$), 소득($b=.26, p<.01; b=.21, p<.05$), 개인정보 보호행동 ($b=-.20, p<.001$)이 있다. 이를 통해 학생이 아닌 사람보다는 학생이 소득이 낮은 집단보다 소득이 높은 집단이 사업자가 빅데이터 환경에서 생산된 개인정보를 활용할 때 더 동의하는 것으로 나타났다. 그러나 개인정보 보호행동 점수가 높을수록 정보의 활용에 동의하지 않는 것으로 나타나 변수들 간의 영향력이 다를 수 있음을 확인할 수 있다.

다중회귀분석 결과를 통해 정보별로, 사업자가 취하는 절차별로 소비자가 빅데이터 환경에서 생산된 개인정보의 활용에 동의하는 정도에 영향을 미치는 요인들이 다를 수 있음을 확인할 수 있다.

[표 4-8] 개인정보 활용에 대한 동의정도의 영향요인

독립변수	종속변수	정보별								절차별							
		공개된 정보				이용내역정보				소극적 절차				적극적 절차			
		b	S.E.	베타	t	b	S.E.	베타	t	b	S.E.	베타	t	b	S.E.	베타	t
상수		2.44	.37		6.53***	2.41	.36		6.63***	2.15	.36		6.00***	2.70	.42		6.43***
성별(여=0)		.13	.08	.08	1.80	.16	.07	.09	2.24*	.16	.07	.09	2.20*	.14	.08	.07	1.67
연령		.01	.00	.11	1.73	.01	.00	.16	2.67**	.01	.00	.19	3.14**	.01	.01	.07	1.17
최종학력(고졸이하=0)		.03	.09	.01	.32	-.08	.09	-.04	-.90	-.03	.09	-.01	-.33	-.02	.10	-.01	-.21
직업(학생 외=0)		.37	.12	.19	3.08**	.31	.12	.16	2.59*	.32	.12	.17	2.75**	.36	.14	.16	2.64**
소득(저=0)	중	.15	.09	.08	1.73	.10	.09	.05	1.13	-.01	.09	-.01	-.12	.26	.10	.13	2.62**
	고	.12	.10	.06	1.27	.14	.09	.08	1.52	.05	.09	.03	.55	.21	.11	.10	1.98*
개인정보 보호행동		-.20	.06	-.16	-3.55***	-.12	.05	-.10	-2.29*	-.11	.05	-.09	-2.07*	-.21	.06	-.15	-3.37**
약관 확인행동 (0=읽지않음)		.09	.09	.04	.95	.04	.09	.02	.41	.01	.09	.01	.14	.11	.10	.05	1.08
유출경험 (없다=0)	있다	-.12	.11	-.06	-1.02	-.10	.11	-.06	-.95	-.08	.11	-.04	-.69	-.14	.13	-.07	-1.14
	모른다	-.08	.12	-.05	-.72	-.07	.12	-.04	-.63	-.10	.11	-.05	-.86	-.06	.13	-.03	-.45
빅데이터 환경에서의 개인정보에 대한 인식		-.00	.06	-.00	-.05	-.12	.06	-.09	-2.01*	-.18	.06	-.14	-3.20**	-.06	.07	.04	.95
의사반영에 대한 인식		.15	.04	.16	3.92***	.13	.04	.15	3.56***	.20	.04	.23	5.51***	.08	.04	.08	1.86
R ² (Adj.R ²)		.09(.07)				.11(.09)				.13(.11)				.10(.07)			
F		4.50***				5.35***				6.21***				4.54***			

유의수준: *** p<.001, ** p<.01, * p<.05

제 5 장 결론 및 제언

제 1 절 요약 및 결론

정보통신 기술이 발전하면서 기존과는 다른 형태의 개인정보, 즉 소비자가 온라인 사이트에서 활동하면서 적극적으로 생산하거나, 기록되는 정보들이 증가하고 있다. 특히 사업자들에 의해 개인정보가 수집·활용되는 과정에서 개인의 프라이버시 침해와 정보 유출로 인한 피해가 발생할 가능성이 존재한다. 따라서 빅데이터 환경에서 생산되는 개인정보에 대해 소비자들이 어떻게 인식하고 있으며, 정보가 사업자에 의해 활용되는 것에 동의하고 있는지 연구할 필요가 있다.

본 연구에서는 빅데이터 개인정보보호 가이드라인에 의거하여 빅데이터 환경에서 생산되는 개인정보의 범위에 속한 개별 정보들을 정리하고, 소비자들이 정보들을 개인정보로 인식하는 수준을 알아보고자 한다. 또한 사업자의 개인정보 활용에 대한 인식과 영향을 미치는 요인에 대해서 살펴보고자 한다. 이러한 문제의식을 바탕으로 진행된 연구의 결과를 요약하여 도출한 결론은 다음과 같다.

첫째, 전반적으로 소비자들이 빅데이터 정보들을 자신의 개인정보라고 높게 인식하였는데, 이를 통해 기존 개인정보의 개념 및 범위를 확장할 필요가 있다. 세부적인 개인정보별 인식의 차이를 살펴보면, 자신이 공개한 이름과 성별 등의 일반적인 인적 사항과 사이트에 접속했을 때 수집되는 실시간 기록을 개인정보로 인식하는 정도가 가장 높았다. 또한 쇼핑 내역과 같은 사이트 거래기록, 사진 및 동영상 등 타 정보에 비해 자신의 개인정보라고 높게 인식하였다. 추후 개인정보와 관련된 논의에 있어서 보다 확대된 시각을 가져야할 필요가 있음을 시사하는 결론이다.

둘째, SNS나 온라인 쇼핑몰 이용률이 높고 일반적인 개인정보에 대한 관심과 우려가 높은 집단이 빅데이터 개인정보에 더 민감하다는 사실을 밝혀내었다. 세부적인 결과를 살펴보면 남성에 비해 여성이, 50대에 비해

다른 연령대가 빅데이터 환경에서의 정보를 개인정보로 더 높게 인식하였다. 또한 개인정보를 보호하기 위한 행동 관련 점수가 높고, 개인정보 유출 경험이 있거나 유출이 되었는지 알 수 없다고 답한 소비자들이 정보들을 자신의 개인정보로 높게 인식하였다. SNS와 인터넷 쇼핑 이용률과 개인정보에 대한 관심, 개인정보 유출경험이 빅데이터 환경에서 생산된 개인정보에 대한 인식에 영향을 미친 것으로 유추할 수 있다.

셋째, 빅데이터 환경에서 이루어지는 사업자의 개인정보 활용에 전반적으로 소비자가 동의하지 않는다는 결과를 고려한다면 빅데이터 개인정보보호 가이드라인에서 명시하고 있는 소비자의 동의 없는 사업자들의 개인정보 활용은 지양되어야 한다. 특히 사업자가 소극적 절차를 통해 개인정보 활용에 대해 고지하였을 경우 동의하지 않는 정도가 높았다. 그러나 사업자가 적극적 절차를 통해 개인정보 활용에 대해 동의를 받았을 경우 소비자들이 활용에 동의하는 정도가 높아졌다. 빅데이터 개인정보보호 가이드라인에서 개인정보를 보호하기 위한 방안으로 강조한 ‘비식별화 조치’ 즉, 사업자가 개인을 알아볼 수 있는 정보를 비식별화한다는 조건을 붙였음에도 동의 정도가 낮다는 것은 소비자가 사업자의 개인정보 활용을 기피하고 있음을 의미한다. 또한 빅데이터 환경에서 생산되는 정보를 활용하여 사업자가 소비자에게 맞춤형 광고와 같은 혜택을 제공할 때에도 동의 정도가 매우 낮았다. 이처럼 빅데이터 환경에서의 개인정보 활용에 대한 소비자들의 동의가 매우 낮은 현 상태에서 빅데이터 산업의 특성상 사전동의를 받는 것이 어렵다는 이유로 개인정보 취급 이후에야 소비자들에게 개인정보 사용에 대한 거부권을 제공하도록 하는 가이드라인은 기업에게 편향된 방안임으로 이를 수정할 필요가 있다.

넷째, 사업자의 개인정보 활용 여부에 대한 선택이 가능할 때 소비자들이 원하는 동의 방식을 살펴본 결과, 소비자들이 더 적극적인 선택권을 발휘하길 선호하였다. 온라인 사이트에 가입한 이후보다 가입하는 당시에 활용여부에 대한 선택권을 발휘하고자 하는 사람들이 더 많았다. 또한 사업자의 개인정보 활용여부를 단 한 번 선택하고 그 선택이 이후 모든 정보 활용에 똑같이 적용되는 방식보다는 정보가 활용될 때마다 선

택권을 가지길 원하였다. 도출된 결과는 소비자들이 개인정보에 대한 선택권을 더 적극적으로 행사하고자 함을 시사한다. 따라서 가이드라인에서 제시한 방안 즉, 소비자에게 사이트 가입 후 단순히 정보 활용에 대한 거부권을 행사할 수 있는 페이지 링크를 제공하는 것보다 사이트 가입 단계에서 더욱 철저한 선택권을 부여해야 한다. 시대의 흐름에 맞추어 사업자의 빅데이터 수집 및 활용을 허용하고 이를 통해 가치를 생산할 수 있도록 독려하는 것이 중요하다. 하지만 앞서 언급한 것과 마찬가지로 가이드라인에서 사전동의 방식을 배제하고 거부권을 제공하는 것은 사업자의 편의에 치우친 것이다. 따라서 정보주체인 소비자들이 개인정보 활용에 동의하지 않고 있는 상황에서 옵트아웃 조치가 적절한지에 대한 재논의가 필요하다.

다섯째, 빅데이터 환경에서 사업자의 개인정보 활용에 대한 소비자의 동의를 높이기 위해서는 활용에 관련된 사실을 명확하게 전달하고, 소비자의 의사를 반영하는 가이드라인이 마련되어야 한다. 다중회귀분석 결과를 살펴보면, 개인정보 보호행동 점수가 높을수록 활용에 동의하지 않는 것으로 나타났다. 또한 빅데이터 환경에서 생산되는 정보를 자신의 개인정보라고 인지할수록 활용에 동의하지 않았다. 하지만 개인정보가 활용되는 과정과 범위에 자신의 의사가 반영될 것이라고 생각하는 소비자가 그렇지 않은 소비자보다 활용에 더 동의하였다. 따라서 소비자의 동의를 바탕으로 개인정보가 활용되기 위해서는 소비자에게 정보 활용에 관련된 사실을 명확하게 전달하여 이에 대한 이해를 넓혀주고, 개인정보를 활용하는 과정과 범위를 확정하는 현시점에서 소비자의 의사를 반영할 수 있는 통로를 제공하는 것이 중요하다.

따라서 빅데이터 개인정보보호 가이드라인을 고안하는 과정에서 빅데이터 환경에서 생산된 개인정보의 활용은 무엇을 의미하며, 개인정보 활용을 통해 생산하는 가치는 무엇인지에 대해 소비자가 납득할 수 있는 소통이 이루어져야 한다. 합당한 절차가 이루어지지 않은 상태에서 가이드라인이 통과된다면 빅데이터 산업의 활성화라는 목표에만 치중될 즉, 개인정보의 ‘보호’가 아닌 ‘활용’에만 치중될 위험이 존재한다.

제 2 절 제언

본 연구는 결론을 통해 다음과 같은 제언을 하고자 한다.

첫째, 소비자정책의 측면에서 소비자의 의견을 반영하여 사업자의 개인정보 활용에 대한 정보주체의 선택권을 강화하는 가이드라인이 수립되어야 한다. 즉, 연구 결과에서 살펴보았듯이 소비자에게 사이트 가입 후 단순히 정보 활용에 대한 거부권을 행사할 수 있는 페이지를 제공하는 것보다 사이트 가입 단계에서 더욱 철저한 선택권을 부여해야 한다. 시대의 흐름에 맞추어 사업자의 빅데이터 수집 및 활용을 허용하고 이를 통해 가치를 생산할 수 있도록 독려하는 것도 중요하지만 현재 정보주체인 소비자들이 개인정보 활용에 동의하지 않고 있는 상황이다. 따라서 빅데이터 환경에서 생산되는 개인정보 활용에 대한 가이드라인이 마련되는 현 시점에서 소비자와 소통하여 의견을 반영할 수 있는 통로를 적극적으로 마련하여야 한다. 사회적 합의가 제대로 이루어지지 않은 상태에서 개인정보 활용에만 초점을 맞춘 가이드라인이 통과된다면 이후 개인정보 유출로 인한 사회적 비용이 막대할 것이다.

또한 개인정보의 보호와 빅데이터 산업의 활성화 사이에 균형을 이루기 위해서는 지속적으로 빅데이터 환경에서 개인정보를 활용하는 사업자를 감시할 수 있는 감독기관을 설립할 필요가 있다. 개인정보를 활용하는 기업이 기술적·관리적 보호조치에 소홀하여 정보 유출이 발생하였을 시 기존보다 강력한 대응이 이루어질 수 있는 제도를 도입할 필요가 있다. 징벌적 손해배상 및 개인정보 유출에 대한 집단소송제도의 도입을 명문화하고 실제로 제도가 활용될 수 있도록 지원하는 것이 방안이 될 수 있다.

둘째, 사업자는 변화하는 정보통신 환경에 발맞추어 빅데이터 산업에 도전하는 현 시점에서 진정한 소비자의 동의를 바탕으로 빅데이터 환경에서 생산되는 정보를 활용하여 가치를 창출하여야 한다. 빅데이터 환경에서 사업자는 개인정보를 수집하고 활용하는 방법과 그 범위에 대해 소비자에게 명확하게 전달하고, 그 내용을 이해시키려는 노력이 중요하다.

또한 개인정보를 보호하기 위해 필요한 안전한 기술적·관리적 보호조치를 도입하고 유지하는데 힘써 소비자의 신뢰를 형성하기 위해 지속적인 노력이 필요하다.

셋째, 소비자는 온라인상에서 자신이 생산하는 정보의 주체임을 자각하고, 정보에 대한 권리를 이해하고 실현하기 위해 노력해야 한다. 계속해서 개인정보의 보호를 외치기만 하고 개인정보에 대해 관심을 갖지 않는다면, 즉 개인정보를 보호하기 위한 행동을 지속적으로 행하지 않는다면 결국 개인정보의 유출로 인한 피해를 입는 것은 개별 소비자이다. 따라서 정부와 기업의 조치만을 기다리고 있는 수동적인 소비자가 아니라 적절한 조치를 요구할 수 있는 조직화된 힘을 보여주는 적극적 주체로 거듭나야 한다.

본 연구의 한계점과 후속연구를 위한 제언은 다음과 같다.

첫째, 빅데이터 환경에서 생산되는 개인정보의 개념과 범위를 확장하기 위해 개인정보가 생산되는 상황과 맥락을 더 세밀하게 분류하는 연구가 필요하다. 본 연구에서는 빅데이터 개인정보보호 가이드라인을 중심으로 개인정보의 종류를 구분하고, SNS에 공개된 정보와 온라인 쇼핑몰에서 발생하는 이용내역정보의 일부를 살펴보았다. 그러나 소비자가 개인정보를 생산하는 상황과 맥락에 따라 개인정보의 분류가 달라질 수 있다. 따라서 유사한 상황과 맥락을 고려하여 정보들을 분류하는 작업이 요구된다.

둘째, 개인정보에 대한 인식을 측정하는 척도를 개발해야 한다. 본 연구는 빅데이터 환경에서 생산되는 개인정보에 대한 인식을 살펴보기 위해 정보를 자신의 개인정보라고 인식하는 수준으로 측정하였다. 그러나 빅데이터의 가장 중요한 특징인 가치에 대한 소비자의 인식을 측정하지 못하였다. 따라서 빅데이터 환경에서의 개인정보가 지닐 수 있는 사회·경제·문화적 가치를 반영하여 이에 대한 소비자의 인식을 측정할 수 있는 보다 엄격한 척도와 변수 개발이 이루어져야 한다.

참 고 문 헌

[국내 문헌]

구혜경·나종연(2014), “소비자의 개인정보 제3자 정보 제공 인식에 관한 탐색적 연구”, *디지털융복합연구*, 12(8), 187-196.

권헌영(2012), *개인정보보호법 시행이후 변화 분석 및 개선방안 연구*, 개인정보보호위원회.

김규남(2014), “빅데이터 2.0시대, 주요 이슈와 정책적 시사점”, *KISDI 프리미엄리포트*, 14(10), 정보통신정책연구원.

김민성·김성태(2014), “개인정보 노출이 잊혀질 권리 수용 의도에 미치는 영향에 관한 연구”, *한국언론학보*, 58(2), 307-336.

김민천(2008), “가상공간에서 수집되는 개인정보의 비교분석”, *사회과학연구*, 24(3), 1-28.

김상우(2012), “EU Data Protection 개정안 이슈 및 시사점”, *정보보호학회지*, 22(6), 39-42.

김선남·이환수(2014), “빅데이터 개인정보보호 가이드라인(안)의 개선방향에 관한 연구”, *정보화정책*, 21(4), 20-39.

김소라·이기춘(2006), “온라인에서의 개인정보유출피해에 대한 위험지각과 개인정보보호수준에 따른 소비자 유형화 및 유형별 관련요인 고찰”, *한국소비자정책교육연구*, 2(2), 45-64.

- 김승한(2013), “빅데이터 시대의 도래에 따른 개인정보보호법의 한계와 개선방안”, *연세 의료·과학기술과 법*, 4(1), 107-145.
- 김정은(2008), “소비자의 개인정보 가치평가와 영향요인에 관한 연구”, 서울대학교 박사학위논문.
- 김종길(2009), “기술위험의 사이버화와 프라이버시권”, *한국사회이론학회*, 35, 245-285.
- 김지운·박지경(2014), “정보사회에서 보건행정 전공 대학생들의 개인정보 보호에 대한 인지 분석”, *디지털융복합연구*, 12(5), 325-334.
- 김혁조(2014), *빅데이터 환경의 개인정보에 대한 소비자 가치추정 연구*, 방송통신위원회.
- 김형길·김정희(2000), “전자상거래에서의 소비자 지각위험에 관한 실증연구”, *경영경제연구*, 2(28), 285-308.
- 김혜선 외(2004), “전자상거래에 있어서 소비자 프라이버시 보호에 관한 연구-인터넷 전자상거래업체의 온라인모니터링조사를 중심으로”, *통상정보연구*, 6(2), 219-242.
- 나종연(2009), *유비쿼터스 환경에서의 개인정보 활용 및 보호방안 연구*, 한국인터넷진흥원.
- 나종연 외(2012), “나에 대해서는 거기까지만!”, 한국소비자학회 통합학술대회, 241-245.
- 나종연(2013), *정보주체 동의권의 실질적 보장을 위한 연구*, 서울대학교

산학협력단.

박정훈·이숙현(2007), “정보 프라이버시와 관련한 개인의 태도 및 행동 경로분석”, *행정논총*, 45(1), 281-307.

성준호(2013), “빅데이터 환경에서 개인정보보호에 관한 법적 검토”, *법학연구*, 21(2), 307-333.

이규철(2013), “신기술(빅데이터) 등장과 개인정보의 보호”, *과학기술법연구*, 19(1), 3-36.

이창범(2012), *개인정보 보호법*, 법문사.

_____ (2013), “개인정보보호법제 관점에서 본 빅데이터의 활용과 보호 방안”, *법학논총*, 37(1), 509-559.

전상민(2015), “개인정보유출에 대한 소비자의 부정적 감정과 행동 반응: Mehrabian과 Russell의 S-O-R 모델을 바탕으로”, *한국심리학회지 소비자·광고*, 16(1), 105-123.

전은정·김학범·염홍열(2012a), “유럽의 개인정보보호 법·제도 동향”, *정보보호학회지*, 22(2), 58-72.

_____ (2012b), “일본·중국의 개인정보보호 법·제도 동향”, *정보보호학회지*, 22(5), 90-105.

정상조·김중보·송옥렬 외(2010), “비식별개인정보의 보호 및 활용에 관한 연구”, *방송통신정책연구*, 4(17), 방송통신위원회.

정영화(2012). “헌법상 정보 프라이버시로서 ‘잊혀질 권리’”, *법학논고*, 39, 567-598.

정원준(2014), “클라우드 컴퓨팅의 활성화를 위한 법적 제문제(I)-개인정보보호 관련 쟁점-”, *정보통신방송정책*, 26(20), 1-13.

조동기·김성우(2003), “인터넷의 일상화와 개인정보보호”, *정보통신정책연구원 이슈리포트*, 11, 1-31.

조성은·유지연(2012), *글로벌 ICT 기업들의 개인정보보호정책 변화에 따른 국내 정책 방향*, 정보통신정책연구원.

정용찬(2013), *빅데이터*, 커뮤니케이션북스.

채은선·신동현(2014), “각국의 개인정보 보호법상 개인정보의 정의 및 해석”, *NIA Privacy Issues*, 13.

방송통신위원회(2014), *빅데이터 개인정보보호 가이드라인*.

방송통신위원회·한국인터넷진흥원(2011), *정보통신서비스 제공자를 위한 개인정보보호 가이드*.

한국방송통신전파진흥원(2014), “유럽의 데이터주도 경제 실현을 위한 움직임 본격화”, *KCA 동향과 전망*, 73-86.

한국인터넷진흥원(2014), “일본 IT종합전략본부, 빅데이터 활용을 위한 개인정보보호법 개정 초안발표”, *Internet & Security Bimonthly*, 3.

한국인터넷진흥원(2014), *2014년 인터넷이용실태조사*.

한은영(2014), “유럽연합(EU), 개인정보보호(Data Protection) 강화 추진”, *정보통신방송정책*, 26(7), 61-70.

행정안전부·한국인터넷진흥원(2011), *사업자를 위한 개인정보보호 질의·응답집*.

개인정보보호법.

정보통신망 이용촉진 및 정보보호에 관한 법률.

[국외 문헌]

Corby, M. J.(2002), “The Case for Privacy,” *Information Systems Security*, 11(2), 9-14.

Cartryse, K. & van der Lubbe, J. C. A.(2003), “Providing Privacy to Agents in an Untrustworthy Environment,” *System*, 38, 1-15.

Cranor, L. F., Reagle, F.J. & Ackerman, M. S.(1999), “Beyond Concern: Understanding Net Users’ Attitudes About Online Privacy,” *AT&T Labs-Research Technical Report*.

EC COM(2014), 442 final, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Towards a Thriving Data-driven Economy.

EU(1995), Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the Protection of Individuals with regard to the Processing of Personal Data and on the Free Movement of such Data.

FTC(2012), *Protecting Consumer Privacy in an Era of Rapid Change: Recommendations for Business and Policymakers*, FTC Report.

ITU(2013). “Big Data: Big Today, Normal Tomorrow,” *ITU-T Technology Watch Report*.

Kelly, J.(2014), “Big Data Vendor Revenue and Market Forecast 2013–2017,” Wikibon.

Manyika, J., Chui, M., Brown, B., et al.(2011), “Big Data: The Next Frontier for Innovation, Competition, and Productivity,” McKinsey & Company.

Oracle(2012), “Big Data for the Enterprise,” *Oracle White Paper*.

Smith, H. J., Milberg, S. J., & Burke, S. J.(1996), “Information Privacy: Measuring Individuals’ Concerns about Organizational Practices,” *MIS Quartely*, 20, 165–195.

Stone, F. F., Gardner, D. G., Gueutal, H. G., & McClure, S.(1983), “A Field Experiment Comparing Information Privacy Values, Beliefs, and Across Several Types of Organizations,” *Journal of Applied Psychology*, 68(3), 459–468.

Tene, O., & Polonetsky, J.(2012), “Big data for all: Privacy and user control in the age of analytics,” *Nw. J. Tech. & Intell. Prop.*, 11, 239-274.

White House(2012), *Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy.*

Youn, S.(2005), “Teenagers’ Perception of Online Privacy and Coping Behavior: A Risk-Benefit Appraisal Approach,” *Journal of Broadcasting & Electronic Media*, 49(1), 86-110.

[기타 자료]

이민형(2015. 1. 15), 빅데이터 개인정보보호 가이드라인, ‘비식별화 조치’ 검증규정 없다, 디지털데일리.

<http://www.ddaily.co.kr/news/article.html?no=126438>

하채림(2014. 6. 19), 개인정보유출 최근 2년8개월간 1억 1천900만 건, 연합뉴스.

<http://news.naver.com/main/read.nhn?mode=LSD&mid=sec&sid1=102&oid=001&aid=0006968721>

허윤(2015. 2. 10), 집단소송제와 징벌적 손해배상제 도입해야, 한국일보.

<http://www.hankookilbo.com/v/0f1fe151aba94d6caab4492f74601de7>

부 록

첨부 1. 연구참여자용 설명서 및 동의서

첨부 2. 설문지

연구참여자용 설명서 및 동의서

연구 과제명 : 빅데이터 환경에서의 개인정보 활용에 대한 소비자인식

연구 책임자명 : 김인혜(서울대학교 생활과학대학 소비자학과 석사과정)

안녕하십니까?

본 연구는 빅데이터 환경에서의 개인정보 및 그 활용에 대한 소비자인식을 이해하고 빅데이터 개인정보보호 가이드라인의 적절한 수정방안을 모색하고자 기획되었습니다. 설문 내용은 빅데이터 환경에서 생산된 개인정보와 그 활용에 대한 인식을 측정하는 문항을 포함하고 있습니다. 본 연구는 개인정보 및 개인정보 활용에 대해 적절히 이해할 수 있는 충분한 능력을 갖추고 온라인 사이트에 가입 경험이 있는 10대~50대 일반인을 연구대상으로 선정하였고, 연구 기간은 2015년 7월 31일까지이며, 응답시간은 약 10분이 소요됩니다. 귀하는 이 연구에 참여하는 데 있어서 직접적인 이득이 없습니다. 그러나 귀하가 제공하는 정보는 개인정보 및 그 활용에 대한 이해를 증진하는 데 도움이 될 것입니다. 귀하의 연구 참여시 감사의 뜻으로 설문을 완료하신 분들에 한해 1000원의 적립금이 지급될 것입니다.

본 연구는 자발적으로 참여 의사를 밝히신 분에 한하여 수행될 것입니다. 귀하께서는 언제든지 어떠한 불이익 없이 참여 도중에 그만 둘 수 있습니다. 만일 응답 도중 귀하가 연구에 참여하는 것을 그만두고 싶으실 때에는 우측 하단의 '중단하기'버튼을 눌러주시기 바랍니다. 본 질문에는 옳고 그른 답이 없으며 수집된 자료는 통계적으로 처리될 뿐 공개되지 않습니다.

본 연구의 개인정보관리책임자는 서울대학교 소비자학과 석사과정 김인혜(02-880-5702)입니다. 본인은 이 연구를 통해 얻은 모든 개인 정보의 비밀 보장을 위해 최선을 다할 것입니다. 이 연구에서 얻어진 개인 정보가 학회지나 학회에 공개될 때 귀하의 개인 정보는 사용되지 않을 것입니다. 그러나 모니터 요원, 점검 요원, 생명윤리심의위원회는 연구 참여자의 개인정보에 대한 비밀 보장을 침해하지 않고 관련 규정이 정하는 범위 안에서 본 연구의 실시 절차와 자료의 신뢰성을 검증하기 위해 연구 결과를 직접 열람할 수 있습니다.

귀하가 본 동의서에 동의 버튼을 누르는 것은 이러한 사항에 대하여 사전에 알고 있었으며 이를 허용한다는 동의로 간주될 것입니다. 본 연구에 대해 질문이 있거나 연구 중간에 문제가 생길 시 연구 책임자인 김인혜(02-880-5702)에게 연락하시고, 만약 어느 때라도 연구 참여자로서 귀하의 권리에 대한 질문이 있으시다면 서울대학교 생명윤리심의위원회(02-880-5153)로 연락하시면 됩니다.

감사합니다.

나는 이 설명서를 읽었으며, 이 연구에 참여하는 것에 대하여 자발적으로 동의합니다. 나는 연구 참여 도중에 언제든지 그만 둘 수 있음을 압니다. 나는 이 연구에서 얻어진 나의 정보에 대한 정보를 현행 법률과 생명윤리심의위원회 규정이 허용하는 범위 내에서 연구자가 수집하고 처리하는데 동의합니다.

동의함

동의하지 않음

첨부 2. 설문지

1. 개인정보와 관련된 귀하의 평소 행동과 일치하는 곳에 표기해 주십시오.

“나는 평소에...”

	전혀 그렇지 않다	별로 그렇지 않다	보통 이다	조금 그렇다	매우 그렇다
인터넷 사이트 가입시 너무 많은 개인정보를 요구하면 다른 사이트로 간다.					
인터넷 사이트 가입을 위해 개인정보를 적을 때, 불필요하다 판단되는 것들에 대해서는 적지 않는다.					
개인정보를 보호하기 위한 방법으로 어떤 것들이 있는지 때때로 찾아보는 편이다.					
인터넷 사이트에 가입할 때 그 사이트에 개인정보보호정책이 있는지 확인한다.					

2. 온라인상에서 개인정보 수집·이용 약관에 동의여부를 클릭하거나 체크해야 하는 경우, 다음 중 귀하의 행동과 일치하는 것은 무엇입니까?

- ① 주로 내용은 읽어보지 않고 체크해야 하는 부분만 찾아서 체크한다.
- ② 대충 스크롤을 내려 보고 체크한다.
- ③ 굵은 글씨나 제목 등 눈에 띄는 부분만 확인하고 체크한다.
- ④ 꼼꼼히 내용을 읽어보고 체크한다.

3. 귀하는 **지난 1년 동안** 개인정보 유출을 경험한 적이 있습니까? (예: 주민등록번호, 이름, 주소 등이 유출된 경험)

- ① 없다 ② 있다 ③ 잘 모르겠다

4. 귀하가 평소에 알고 있는 사실과 일치하는 곳에 표기해 주십시오.

“온라인 사이트 가입 단계에서...”

그렇다 아니다 모른다

내가 필수적으로 동의한 개인정보를 사업자가 수집하여 활용할 수 있다.			
내가 선택적으로 동의한 개인정보를 사업자가 수집하여 활용할 수 있다.			
내가 제3자에게 제공하는 것을 동의한 개인정보를 사업자가 수집하여 활용할 수 있다.			

5. 빅데이터 환경에서의 개인정보와 관련된 귀하의 평소 생각과 일치하는 곳에 표기해 주십시오.

“나는 _____를 나의 개인정보라고 생각한다.”

		전혀 그렇지 않다	별로 그렇지 않다	보통 이다	조금 그렇다	매우 그렇다
자 신 이 전 체 공 개 한	일반 인적 사항 (이름, 성별, 학력, 직업 등)					
	온라인 인적 사항 (아이디, 이메일 등)					
	글					
	사진 및 동영상					
	위치정보					
타 인 이 전 체 공 개 한	사회적 정보 (SNS 인맥, 동호회 등)					
	내가 포함된 글					
	내가 포함된 사진 및 동영상					
	내가 포함된 위치정보					
사 이 트	내가 포함된 사회적 정보 (SNS 인맥, 동호회 등)					
	사이트 접속기록 (접속 시간, 장소 등)					
	사이트 접속 실시간 기록 (현재 위치 정보 등)					
	사이트 클릭기록 (광고, 방문 사이트 등)					
	사이트 검색기록 (검색어 등)					
사이트 거래기록 (쇼핑 내역 등)						

빅데이터 환경에서 생산된 개인정보에는 **공개된 정보와 이용내역정보** 등이 있습니다.

공개된 정보 예시 : 전체 공개한 사진, 글, 동영상, 인맥 정보 등
 이용내역정보 예시: 서비스 이용기록, 접속정보, 거래기록 등

사업자의 개인정보 **활용**에는 개인정보의 **수집·저장·조합·분석 및 제3자 제공** 등이 포함됩니다.

6. 빅데이터 환경에서 생산된 개인정보 활용과 관련된 귀하의 평소 생각과 일치하는 곳에 표기해 주십시오.

항목	전혀 그렇지 않다	별로 그렇지 않다	보통 이다	조금 그렇다	매우 그렇다
빅데이터 개인정보가 활용되는 과정 에 내 의사가 충분히 반영될 것이다.					
빅데이터 개인정보가 활용되는 범위 에 내 의사가 충분히 반영될 것이다.					

7. 귀하의 평소 생각과 일치하는 곳에 표기해 주십시오.

“사업자가 나의 개인정보를 모두
내가 누군지 알 수 없게 처리한 상태입니다.”

		전혀 그렇지 않다	별로 그렇지 않다	보통 이다	조금 그렇다	매우 그렇다
공 개 된 정 보 를	내가 모두에게 공개한 것이므로 별다른 조치없이 사업자가 활용해도 된다.					
	사업자가 활용한다는 것을 쉽게 확인할 수 있으면 , 사업자가 활용해도 된다.					
	사업자가 활용할 때 나에게 알리면 활용해도 된다.					
	사업자가 활용할 때 나에게 동의를 받는다면 활용해도 된다.					
이 용 내 역 정 보 를	온라인 사이트에 가입하는 것만으로도 사업자가 활용해도 된다.					
	사업자가 활용한다는 것을 쉽게 확인할 수 있으면 , 사업자가 활용해도 된다.					
	사업자가 활용할 때 나에게 알리면 활용해도 된다.					
	사업자가 활용할 때 나에게 동의를 받는다면 활용해도 된다.					

8. 온라인 사이트 가입 시 **빅데이터** 환경에서 생산된 개인정보 활용에 **동의하지 않을 경우** 해당 사이트에 **가입을 할 수 없다면**, 사업자의 개인정보 활용에 동의할 것입니까?

- ①무조건 동의한다 ②가입을 고려(유보)하겠다 ③무조건 가입하지 않는다

9. 온라인 사이트 가입 시 **빅데이터** 환경에서 생산된 개인정보 활용에 **동의하지 않아도** 해당 사이트의 **가입을 할 수 있다면**, 아래 중 어느 방식을 선호하십니까?

항목	선호방식
온라인 사이트에 <u>가입할 때</u> 사업자의 개인정보 활용에 대한 선택을 할 수 있으면 좋겠다.	
온라인 사이트에 <u>가입한 후</u> 사업자의 개인정보 활용에 대한 선택을 할 수 있으면 좋겠다.	
항목	선호방식
사업자가 개인정보를 활용할 때 <u>단 한 번으로</u> 사업자의 개인정보 활용에 대한 선택을 할 수 있으면 좋겠다.	
사업자가 개인정보를 <u>활용할 때마다</u> 매번 사업자의 개인정보 활용에 대한 선택을 할 수 있으면 좋겠다.	

10. 아래의 문항들 중 귀하의 생각과 가장 일치하는 칸에 표시해주십시오.

“사업자가 빅데이터 환경에서 수집된 개인정보를 활용해서...”

	전혀 그렇지 않다	별로 그렇지 않다	보통 이다	조금 그렇다	매우 그렇다
나에게 메일, 문자, 전화 등을 통해 광고를 하는 것은 괜찮다.					
나에게 맞춤형 광고를 제공하는 것은 괜찮다.					

<다음은 귀하의 **일반적 사항**에 관한 질문입니다.>

1. 귀하의 **성별**은 무엇입니까?

- ①남자 ②여자

2. 귀하의 **연령**은 어떻게 되십니까? 만 _____ 세

3. 귀하의 **최종 학력**은 무엇입니까?

- ① 중졸 이하
② 고졸 이하(중퇴 포함)
③ 대졸/전문대졸(중퇴 및 재학 포함)
④ 대학원 이상(재학 포함)

4. 귀하의 **직업**은 무엇입니까?

- ① 기업인, 경영직 ② 사무, 기술직 ③ 판매, 서비스직 ④ 자영업
⑤ 생산, 기능, 노무직 ⑥ 학생 ⑦ 주부 ⑧ 기타(_____)

5. 귀하가 속한 **가계(현재 함께 살고 있는 가구 기준)의 월 소득**은 어느 정도
입니까?

- ① 100만원 미만
② 100만원 이상 - 300만원 미만
③ 300만원 이상 - 500만원 미만
④ 500만원 이상 - 700만원 미만
⑤ 700만원 이상

Abstract

Consumer Perception on Using Personal Information in Big Data Environment

Inhye, Kim

Dept. of Consumer Science

The Graduate School

Seoul National University

As the information communication technology has developed, other forms of personal information including photos, posts and online behavior which are produced by active consumers in online are increasing. In particular, there is a possibility that invasion of privacy occurs during the process of operators' collecting and using such information. Therefore, it is necessary to study which information consumers perceive as their own personal information and how they think about the usage of information by operators in big data environment. In this study, we reveal the consumers' perception on personal information and the usage of it in big data environment.

This study carried out a survey targeting 574 consumers around Korea in their teens, 20s, 30s, 40s and 50s, used panel by an on-line survey company. The main conclusions of the study are as follows.

First, consumers highly perceive all the information produced in big data environment as their personal information. This result leads to the conclusion that we need to expand the scope of study on personal information. Also, consumers frequently using online shopping mall and SNS are more sensitive about personal information created in big data environment.

Second, most consumers do not give consent to the operators' usage of personal information. This result reveals that usage of personal information produced in big data environment should be careful. While consumers do not accept the usage when operators do not inform on it well, they consent a little more when operators inform on it and get consumers' consent directly.

Third, consumers want to play an active role when choosing whether operators get the right to use personal information or not. They prefer showing acceptance or rejection on the usage of personal information when they join websites. Also, they want to get noticed every time operators use personal information.

Fourth, consumers give consent to operators' using personal information when they perceive that their opinion is reflected throughout the process of usage. That is why we need to hear the voice of consumers when amending Big data Guideline in Korea.

Those are several suggestion that I draw from the result.

First, Big Data Guideline should be amended reflecting consumers opinion. As previously stated, consumers want to play an active role about operators' usage of personal information. Therefore, we need to give consumers the right to choose whether their personal information

would be used or not at the first stage of joining websites.

Second, operators should create value in a new era of Big Data based on the consumers' consent when collecting and using personal information. Recently almost every consumer has suffered with serious personal information leakage and wants to avoid further problem. Therefore, operators should protect personal information thoroughly and manage it properly.

Third, consumers need to realize that they are owners of their personal information. They should pay attention to the process of the usage of information and speak up for the right that their personal information needs to be protected by operators.

Here are limits of study and suggestions for a follow-up study.

In this study, personal information is classified by Big Data Guideline in Korea. However, personal information could be categorized differently depending on situations. Therefore, it is recommended that information should be researched based on specific situation. Also, we need to develop elaborate measurement of social, economic and cultural value of personal information.

**keywords : Big Data Guideline in Korea,
personal information in Big Data environment,
perception on personal information,
personal information usage,
giving consent to the usage of personal information**

Student Number : 2013-21527