

최신 ICT 이슈

I. 페이스북 사태의 배후, 사람보다 정확한 컴퓨터의 성격 판단 기술

소비자 분석 기술 제공기업인 케임브리지 애널리티카가 페이스북 앱을 이용해 수집한 5,000만 명 이상의 개인정보를 분석하여 타깃층에 가짜뉴스를 유포하는 등 세계 각국의 선거전에 개입했다는 의혹이 제기되고 있는 가운데, 이러한 과정을 허용한 페이스북에 대해 강력한 사회적 비판이 가해지고 있음. 페이스북에 최대 위기를 가져 온 이번 사태의 배경에는 ‘좋아요’ 정보 등을 기계학습으로 분석해 사람의 성격을 정확히 판정해 내는 기술이 자리하고 있음

- 페이스북은 최근 사용자 개인정보를 부정하게 관리했다는 강한 의혹을 받고 있으며, 특히 미 대선에 영향을 미쳤을 가능성 때문에 정보 관리 책임을 엄격히 추궁당하고 있음
 - ▶ 페이스북이 맞이라고 있는 창사 이래 최대의 위기는, 미 대선 당시 트럼프 캠프의 캠페인을 담당했던 데이터 분석 기업 “케임브리지 애널리티카(Cambridge Analytica)”가 페이스북 계정 5,000만 개의 개인정보를 비밀리에 입수하여 활용했다는 보도로부터 시작되었음
 - ▶ 보도에 따르면, 케임브리지 대학의 알렉산드르 코건 교수는 2014년에 “디시이즈유어디지털라이프(thisisyourdigitallife)”라는 페이스북 앱을 만들었는데, 간단한 심리 테스트에 참여하여 개인정보와 심리상태의 정보를 제공하면 돈을 받을 수 있는 앱이었음
 - ▶ 이 앱은 사용자뿐만 아니라 사용자의 친구 데이터까지 수집할 수 있었다고 하며, 코건 교수는 앱을 통해 모든 고객 정보를 상업적 목적으로 활용할 수 없다는 페이스북 규정을 어기고 심리 테스트 결과로 획득한 데이



〈자료〉 The Rush Limbaugh Show

[그림 1] 페이스북 고객 데이터 무단 접근 사태

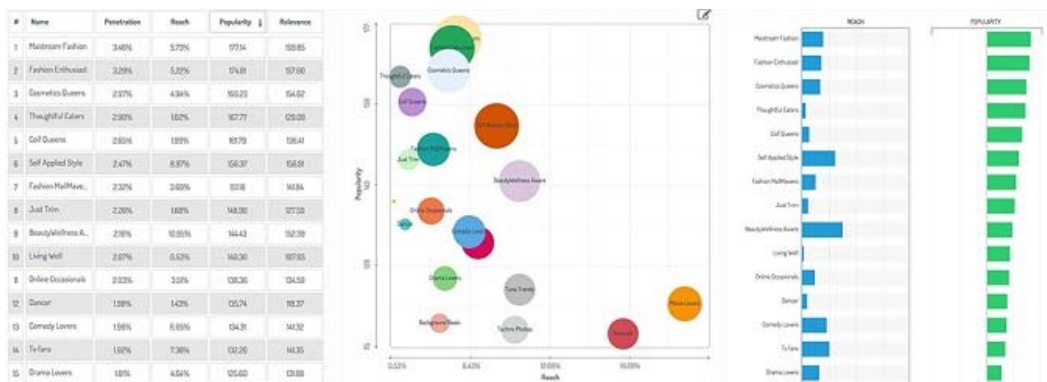
* 본 내용과 관련된 사항은 산업분석팀(☎ 042-612-8296)과 최신ICT동향 컬럼리스트 박종훈 집필위원(soma0722@naver.com ☎ 02-576-2600)에게 문의하시기 바랍니다.

터를 케임브리지 애널리티카에 판매하였음

- ▶ 앱 사용자는 약 27만 명이었으나 실제로 코간 교수가 데이터를 획득한 계정은 8,700만 개였고, 이를 사들인 케임브리지 애널리티카가 데이터를 분석하여 고도로 표적화된 선거 캠페인을 전개했다는 것이 의혹의 골자임
- ▶ 페이스북이 2015년에 데이터의 삭제 요청을 했음에도 불구하고 2017년까지 여전히 삭제되지 않았다고 하는데, 이는 2016년 대선에서 이 데이터들의 분석 결과가 확실히 사용되었음을 의미함
- ▶ 보도 직후 파장이 커지자 페이스북은 자사 플랫폼 내에서 사용자 정보의 유용을 막기 위한 6가지 대책 방안을 발표하며 무마에 나섰지만, 대책 방안의 내용이 그 동안 데이터 관리가 얼마나 허술했는지를 반증하는 것이어서 사용자 반발이 사그러들지 않고 있음
- ▶ '#deletefacebook'이라는 해시태그와 함께 페이스북 계정 삭제 운동이 이어지고 있으며, 실제 일론 머스크가 테슬라와 스페이스X의 페이스북 페이지를 모두 삭제하는 등 ICT 업계 내부에서도 강한 비판이 쏟아지고 있음

■ 논란의 출발점이 된 케임브리지 애널리티카는 런던 소재 스타트업으로 데이터 사이언스 기법으로 소비자와 유권자의 성향을 파악하는 “심지(心誌) 분석” 기술을 개발하였음

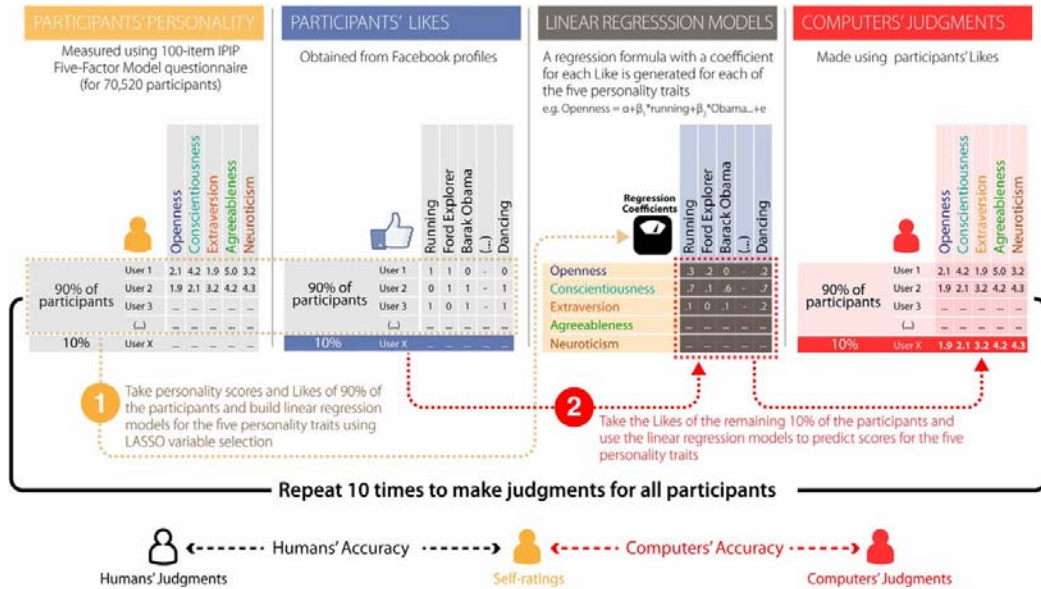
- ▶ 케임브리지 애널리티카는 현재 두 개의 솔루션을 제공하고 있는데, 광고 마케팅 기업에게는 소비자를 대상으로 한 타겟 광고 도구를, 선거 관계자들에게는 유권자 성향을 분석하는 선거 도구를 제공하고 있음



〈자료〉 Cubeyou

[그림 2] 사이코그래프의 예

- ▶ 이 기업은 페이스북의 개인정보를 사들여 유권자의 정치 성향을 파악한 것으로 의심받고 있는데, 소비자와 유권자를 분석할 때 “Psychographic Analysis(심지 분석, 사이코그래프)”라는 기술을 사용한다고 함
 - ▶ 사이코그래프(Psychograph)는 개성의 각 측면을 나타내는 몇 가지 항목을 선정하여 특정 개인에 관해 그 항목마다 특성을 평가하여 그래프로 나타낸 것으로 ‘심지(心誌)’ 또는 “인격 프로파일”이라 번역됨
 - ▶ 케임브리지 애널리티카는 페이스북의 프로필 정보를 심지 분석 기술을 사용해 분석함으로써 각 이용자의 성격 특성을 도출했으며, 특히 이용자의 ‘Like(좋아요)’ 버튼 클릭 정보를 토대로 성격을 파악할 수 있다고 함
- **심지 분석 모델에 따르면 ‘좋아요’를 누르는 패턴과 개인의 성격 사이에 강한 상관관계가 있다고 하며, 이를 근거로 공화당과 민주당 지지자를 비교적 정확히 구분할 수 있다고 함**
- ▶ 케임브리지 애널리티카가 사용한 심지 분석 알고리즘은 화가 살바도르 달리(Salvador Dali)를 좋아하는 사람은 개방적인 성격으로, 조깅을 취미로 하는 사람은 꼼꼼한 성격으로 판정한다고 함
 - ▶ 또한, 애니메이션이나 만화를 좋아하는 사람은 사교성이 없다고 진단하는데, 이러한 개성 분류 정보는 선거에 응용할 수 있는 다양한 인사이트를 제공할 수 있음
 - ▶ 이 심지 분석 모델은 공화당 지지자와 민주당 지지자를 정확하게 판정할 수 있고, 공화당 지지자 중에서도 “폐쇄적이며 사소한 일에 걱정이 많은 성격”의 유권자를 식별함
 - ▶ 알고리즘은 이 폐쇄적이고 기우가 심한 사람들의 그룹을 트럼프의 핵심 지지층으로 알려진 저학력 고령의 남성 공화당 지지자로 추정한다고 함
- **케임브리지 애널리티카의 심지 분석 알고리즘은 “컴퓨터 기반의 개성 판단” 기법에 이론적 근거를 두고 있으며, 컴퓨터의 판단이 사람의 판단보다 정확하다고 보고 있음**
- ▶ 심지 분석 기법의 기반이 되는 이론은 케임브리지 대학 심리학과와 스탠퍼드 대학 컴퓨터 공학과 연구팀이 공동으로 개발하였음
 - ▶ 이 이론은 이용자의 성격이 5가지 요소, 즉 Openness(개방성), Conscientiousness(양심), Extraversion(외향성), Agreeableness(어울림), Neuroticism(노이로제)로 구성되어 있다고 보며, 이들 요소가 어떤 비중으로 구성되느냐에 따라 사람의 성격이 결정된다고 봄
 - ▶ 양 대학은 사이코그래프 분석에 대한 논문 “Computer-based personality judgments are more accurate than those made by humans(컴퓨터 기반의 개성 판단이 사람의



〈자료〉 Michal Kosinski et al.

[그림 3] 페이스북 “좋아요 누르기” 정보로 성격을 파악하는 심지 분석 기법

판단보다 정확하다”)를 통해 자신들이 연구한 기술 내용에 대해 발표하였음

- ▶ 이들은 방법은 두 단계로 이루어지는데, 우선 70,520명의 피험자를 대상으로 Personality Test(성격 진단 테스트)를 실행해 성격을 판정하며, 성격은 앞서 언급한 5가지 요소로 구성되고, Personality Test에 의해 5개 요소의 가중치가 결정됨
- ▶ 다음으로 피험자의 페이스북 개인 프로필 정보를 참조하는데, ‘좋아요’ 버튼을 누른 대상이 무엇인지(예; Running, Ford Explorer, Barak Obama 등)를 파악하여 피험자가 어떤 항목에 관심을 나타내고 있는지를 식별하려 하였음
- ▶ 연구팀은 필요한 개인 정보를 수집하기 위해 ‘myPersonality’라는 앱을 개발했는데, 이용자가 이 앱으로 Personality Test를 해볼 수 있게 하는 대신, 이용자의 허락 하에 또한 학술 연구만을 위해 이용된다는 조건 하에 ‘좋아요’를 눌렀을 때 그 정보가 수집되도록 하였음
- ▶ Personality Test와 ‘좋아요’ 클릭 정보가 모이고 나면, 데이터 간의 관련성을 선형회귀분석(Linear Regression) 기법으로 도출하였음
- ▶ 성격과 ‘좋아요’ 클릭 사이의 관련성을 정의하는 변수를 도출한 것인데, 가령 외향성이 강한 사람은 Running, Ford Explorer, Barak Obama 등의 항목을 각각 어떤 패턴으로 좋아하는지를 추정하였음

- ▶ 이런 과정을 통해 나온 결정 모델을 이용하여 실제 판정을 실시했으며, Personality Test를 받지 않은 피험자라도, '좋아요'를 누른 정보를 이 모델에 입력하기만 하면 개인의 성격을 판정하도록 하였음
 - ▶ 판정 결과는 위의 5 가지 구성 요소가 각각 어떤 비율인지를 추정하는데, 이 모델은 '좋아요' 클릭 정보만으로 그 사람의 성격을 추정할 수 있음을 제시하였고, 그 판정 결과가 사람이 직접 성격을 판정하는 것보다 정확하다고 주장한 점에서 주목을 받았음
- **케임브리지 애널리티카는 미국 대선을 앞두고 새로운 선거 캠페인 모델을 개발하고자 했으며, 두 대학의 심지 분석 기술과 유사한 기법으로 모델을 개발하였음**
- ▶ 당초 케임브리지 애널리티카는 '좋아요' 심지 분석 기술을 개발한 케임브리지 대학 연구팀에 협조를 요청했으나 이루어지지 못했고, 대신 이 연구에 정통한 같은 대학의 알렉산드르 코건 교수에 도움을 요청했고, 코건은 동일한 방법에 기반한 모델을 개발하였음
 - ▶ 코건 교수는 위에서 소개한 두 대학 연구팀의 'myPersonality' 앱을 모방한 성격 진단 테스트 앱으로 'thisisyourdigitallife'를 개발했으며, 페이스북 이용자 27만 명이 이를 이용했음
 - ▶ 이용자는 이 앱을 통해 자신의 성격을 알 수 있었으며 소정의 돈도 받았으나, 동시에 앱이 개인정보에 접근하는 것을 허용해 프로필 데이터가 수집되도록 하였음
 - ▶ 또한, 이 앱이 사용자의 친구의 프로필 정보에 접근하도록 요청했기 때문에, 코건 교수는 5,000만 명 분의 개인정보를 입수할 수 있었고, 이들 데이터를 대상으로 심지 분석을 수행하였으며 3,000만 명의 성격을 추정하였음
 - ▶ 코건 교수는 이렇게 분석한 정보를 케임브리지 애널리티카에 판매한 것인데, 2014년 당시 페이스북은 이용자의 허가를 획득하면 제삼자가 개인 정보를 수집하는 행위는 인정했으나, 그렇게 수집한 정보를 다시 다른 사람에게 넘기는 것은 금지하고 있었음
 - ▶ 이 부분이 문제의 핵심으로 결국 케임브리지 애널리티카가 페이스북의 규정을 위반하며 개인 정보를 부정하게 입수한 것으로 보이며, 비록 자신들은 몰랐다고 부정하고 있지만 애초 코건 교수에게 도움을 요청해 시작된 일이기 때문에, 영국 정부는 데이터 남용 혐의로 수사에 착수했음
- **케임브리지 애널리티카에 전달된 개인 정보가 어떤 식으로 사용되었는지에 대한 정확한 조사 결과는 아직 발표되지 않았지만, '가짜뉴스' 발송에 적용했다는 내부 증언이 나오고 있음**
- ▶ 선거용 분석 도구를 제공하고 있는 기업이기 때문에 사이코그래프를 선거전에 적용하려 했음

은 불문가지인데, ‘좋아요’ 누르기 정보를 통해 유권자의 성격과 정치적 성향을 파악할 수 있다면 고도로 타게팅이 가능한 최적의 캠페인을 전개할 수 있기 때문

- ▶ 케임브리지 애널리티카의 전 직원 크리스 와일리가 영국 의회 청문회에 출석하여 이 분석 모델을 미국 대통령 선거에 어떻게 적용했는지에 대해 증언한 내용은 가히 충격적인데, 일반적인 선거 캠페인의 목적을 넘어서 사용했음을 폭로하고 있음



〈자료〉 CNN

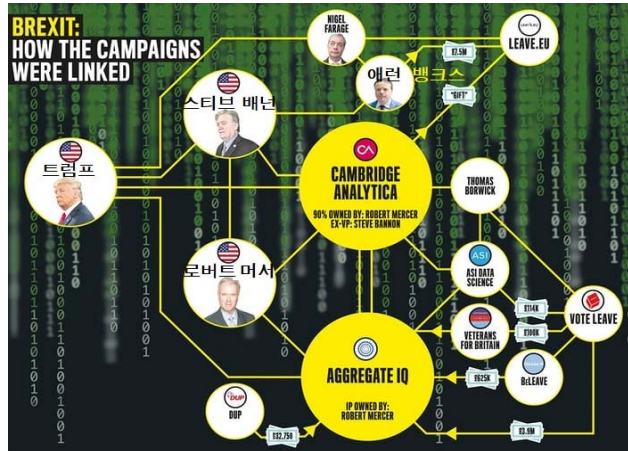
[그림 4] 전 직원 크리스 와일리의 폭로

- ▶ 증언에 따르면 케임브리지 애널리티카는 유권자의 심리적, 정신적인 취약점을 찾아낼 목적으로 이 분석 모델을 사용했으며, 또한 그 약점을 자극하는 가짜 뉴스를 타깃 전송함으로써 유권자를 특정 방향으로 치우치게 하고 트럼프 후보에 투표하도록 독려했다고 함
 - ▶ 크리스 와일리 본인은 이 심지 분석 모델을 운영하는 과정에는 관여하지 않았으며 실제 어떤 식으로 전개되었는지는 알 수 없다고 말해, 이 부분은 정확한 진상 조사 결과가 나와야 확인할 수 있지만 증언에 의하면 불법 선거운동에 활용된 것이라는 말이 됨
 - ▶ 이런 의혹에 대해 “심지 분석”은 특별한 것이 아니고, 소비자의 특성을 파악하여 최적의 광고 메시지를 전달할 수 있다는 가능성 때문에 이미 타깃 광고에서 많이 사용되고 있으며, 넷플릭스 등은 시청자가 선호하는 영화 추천에도 이 모델을 사용한다는 반론도 있음
 - ▶ 또한, 선거 전문가들은 유권자의 마음을 움직이는 것은 어려운 일이기 때문에 심지 분석 기술이 유권자에게 얼마나 영향을 미치는지는 의문이며, 케임브리지 애널리티카가 대통령 선거에 실제 미친 영향은 제한적이었을 것이라는 견해를 내놓고 있음
 - ▶ 그러나 실제 목적을 달성하지 못했거나 영향력이 작다고 해서 불법적인 시도를 용인할 수는 없는 것이며, 심지 분석을 페이스북의 최대 폐해 중 하나로 지적되고 있는 ‘가짜뉴스’ 유통에 이용하려 했다는 점 만으로 케임브리지 애널리티카는 비난을 면치 못하고 있음
- 케임브리지 애널리티카에 대한 비난은 미국뿐 아니라 영국, 프랑스 등 다른 나라에서도 거세게 일고 있으며, 주로 극우파에 유권자 분석 도구를 제공했다는 의혹을 사고 있음
 - ▶ 크리스 와일리의 폭로 이후 영국 고등법원은 케임브리지 애널리티카에 대한 수색 영장을 발부

해 압수 수색을 했는데, 이는 케임브리지 애널리티카가 2016년 브렉시트(영국의 EU 탈퇴) 국민투표 여론전에도 관여했다는 추가 폭로가 나왔기 때문

- ▶ 추가 폭로의 근거로 제시된 것은 케임브리지 애널리티카가 작성한 “EU 탈퇴 국민투표를 위한 빅데이터 솔루션”이라는 제목의 문서인데, 유권자, 정치인, 언론인들 중에 브렉시트 찬성자가 누군지 구분해 낼 수 있다는 내용을 담고 있음

- ▶ 케임브리지 애널리티카는 즉각 브렉시트 투표 과정에서 부정 한 일을 하지 않았다고 부인했지만, 투표 결과가 예상 외로 탈퇴 찬성이 4% 포인트 더 높게 나왔기 때문에, 탈퇴 반대 진영에서는 가짜뉴스 등에 의해 여론이 호도되었을 가능성을 강하게 의심하고 있음



〈자료〉 The Guardian

[그림 5] 미 대선과 브렉시트 선거의 연결

것으로 알려지고 있으며, 2014년 4월 초에는 나이지리아 정부가 2015년 자국 선거에 개입한 혐의로 케임브리지 애널리티카에 대한 조사를 시작하였음

- ▶ 케임브리지 애널리티카가 미국 대선을 비롯하여 여러 나라의 선거에서 극우파의 여론전을 지원한 것 아니냐는 의혹이 나오고 있는 배경에는 이 기업과 관련된 인사들이 주로 극우적 성향을 가지고 있고 서로 연관이 되어 있기 때문임
- ▶ 케임브리지 애널리티카에는 트럼프 대통령의 후원자이자 헤지펀드 억만장자인 로버트 머서가 1,500만 달러를 투자했는데, 이는 이 기업 지분의 90%에 해당하는 것임
- ▶ 한편 전 백악관 고문으로 트럼프의 오른팔로 불렸던 스티브 배넌 역시 케임브리지 애널리티카의 부사장을 맡은 바 있으며, 이 기업에 의해 벌어진 여론 조작 과정의 배후로 지목되고 있음
- ▶ 스티브 배넌은 영국의 브렉시트를 주도한 애런 벅크스와 관련을 맺고 있는 것으로 알려져 있으며, 브렉시트 선거를 앞두고 벌어진 각종 토론회에서 케임브리지 애널리티카는 애런 벅크스와 자리를 함께 한 바 있음
- ▶ 이런 배경 때문에 케임브리지 애널리티카 스캔들은 단순히 페이스북의 이용자 정보를 불법적

으로 획득하여 선거 캠페인에 이용했다는 차원의 문제가 아니라, 극우 정치세력의 여론 조작이라는 의혹을 받고 있으며, 진상 조사 결과에 국제적 관심이 쏠리고 있는 것임

■ 한편, 케임브리지 애널리티카의 불법행위 유무와 관계없이 페이스북은 이번 사태의 본질적인 원인 제공자로 개인 데이터 관리의 책임을 엄격하게 추궁당하고 있음

- ▶ 사태가 터지고 난 후 페이스북은 정보 유출 사실을 부인하다가 문제가 커지자 유출 사실을 인정하고 조치를 취하겠다고 발표함으로써 화를 키웠다는 비판을 받고 있음
- ▶ 페이스북은 현재 일련의 개인정보보호 대응을 진행하고 있으며, 지금까지 20개 화면에 분산되어 있던 개인정보 설정을 하나의 화면으로 통합하여 프로필 설정 방식을 알기 쉽게 했고 정보 관리를 용이하게 하였음
- ▶ 또한, 제삼자가 생성하는 분석 데이터의 제공도 중단시켰는데, 데이터 분석 기업인 Experian (익스피리언)과 Acxiom(액시엄) 등은 오프라인 데이터를 분석하여 이를 광고주에게 제공하고 있었는데, 이를 중단한다고 발표하였음
- ▶ 이번 사태에서 이용자들이 가장 크게 충격을 받은 것은 페이스북의 서드파티 앱이 사용자 프로필뿐 아니라 그 친구들의 정보까지 접근할 수 있었다는 것이고, 놀랍게도 이런 정보 수집은 당시 페이스북의 서드파티 앱 이용 정책 하에 허용되는 행위였다는 사실임
- ▶ 또한, 플랫폼 운영업체들이 그 어느 때보다도 데이터 보호의 중요성을 외치고 있음에도 불구하고, 평균적인 데이터 유출 사건의 규모와 영향력은 점점 커지고 있으며, 5,000만 명의 데이터가 유출된 이번 페이스북 사태는 이러한 경향을 다시 한번 확인해 주었음

■ 사실 이런 문제는 페이스북을 포함하여 모든 플랫폼 업체가 안고 있는 근본적인 딜레마에서 비롯되기 때문에 어쩔 수 없는 측면이 있고 또한 근원적 해결도 쉽지 않음

- ▶ 모든 플랫폼들은 사용자들의 데이터를 보호해야 한다는 의무와 이 데이터를 이용해 돈을 벌고자 하는 욕망 사이의 충돌에서 외줄타기를 할 수밖에 없음
- ▶ 페이스북을 비롯한 플랫폼들의 자산은 이용자 기반, 다시 말해 사용자 데이터이며 플랫폼은 이 데이터를 서드파티 개발자들에게 어떻게든 홍보를 해야만 하는 입장에 있음
- ▶ 사용자 데이터를 서드파티 앱에서 사용하도록 제공한다면, 더 많은 개발자들이 자신들의 플랫폼을 이용하게 될 것이고, 플랫폼의 영향력은 더욱 커질 것이며, 개별 사용자들의 성향에 맞춘 광고를 게재함으로써 더 높은 수익을 얻을 수 있기 때문임
- ▶ 페이스북이 이용자들의 비난이 원체 거세자 우선 모면을 위해 서드파티 앱의 접속을 제한하고

있지만, 한 컨에서는 이런 조치가 기업 마케터들과 개발자에게 치명적이며 페이스북의 마케팅 가치를 하락시킬 것이라는 전망이 나오는 것은 이런 이유에서임

- ▶ 따라서 이번 페이스북의 조치들을 통해 사용자들의 데이터가 안전하게 보호될 수 있을 것이란 기대는 선불리 할 수 없으며, 사용자의 데이터를 보다 교묘하게 이용할 수 있는 새로운 방법이 마련될 것이라고 보는 것이 보다 현실에 가까운 전망일 것임

■ 그러나 페이스북 사태의 파장이 원체 큰 탓에 당분간은 GDPR 등 데이터 보호가 강조될 것으로 전망되며, 반면 “개인 데이터=돈”이라는 사실도 보다 명확해질 것으로 보임

- ▶ 대규모 개인 데이터 유출 사태가 터지자 소비자들의 사생활이 보호받을 권리에 관해 명쾌한 기준을 제시하고 있는 유럽연합의 GDPR(General Data Protection Regulation, 보편적 데이터 보호 규제) 등이 해결책으로 제시되고 있음
- ▶ GDPR은 2018년 5월 25일부터 시행되는 EU의 정보보호 규약으로, 규제 내용 가운데 가장 널리 알려진 것은 “잊힐 권리(the right to be forgotten)”임
- ▶ EU와 관계를 맺고 있는 전 세계 모든 기업은 EU의 사용자가 자신의 데이터를 삭제해 달라고 요구하면 이에 응해야 하며, 사용자는 자신의 데이터를 “널리 이용되는 기기에 의해 판독이 가능한” 형식으로 인도받을 수도 있음
- ▶ 페이스북의 서드파티 앱 제한이 지속되고, GDPR 등이 강화된다면, 현재 사용하는 앱이 제대로 작동하지 않을 것이기 때문에 마케터들과 개발자들은 물론 이용자들의 불만의 목소리도 나올 가능성이 있음
- ▶ 그러나 데이터 보호 전문가들은 불편함을 감수하고서라도 사용자들이 제로 트러스트(zero trust), 즉 모든 정보에 대한 접속을 차단한 상태에서 앱의 정보 수집 메커니즘을 정확히 살펴 보고, 안전하고 필요한 경우에만 접속을 허용하는 훈련을 해야 한다고 조언하고 있음
- ▶ 한편, GDPR과 같은 규정이 강화되고, 제로 트러스트에서 시작하는 소비자들이 많아진다면, 소비자들의 데이터가 곧 돈이라는 것이 보다 명확해지며, 돈으로 개인 데이터를 사는 것이 보편화될 가능성도 있음
- ▶ 지금은 익명의 빅데이터 분석이라는 논리 하에 불특정 다수의 데이터가 수집되는 경우도 많지만, 이런 경우에도 비용을 지급하는 관행이 마련되고, 비용 지급을 중개하는 에이전트들이 출현하는 것도 예상해 볼 수 있음

[참고문헌]

- [1] New York Times, “‘You Are the Product’: Targeted by Cambridge Analytica on Facebook,” 2018. 4. 8.
- [2] Bloomberg, “The Problem Is Facebook, Not Cambridge Analytica,” 2018. 3. 20.
- [3] The Rush Limbaugh Show, “Cambridge Analytica Did Nothing Unusual,” 2018. 3. 19.
- [4] The Guardian, “The great British Brexit robbery: how our democracy was hijacked,” 2017. 3. 7.