

Hanbit eBook

Realtime 19

빅데이터 시대의 개인정보보호와 사생활

프라이버시와 빅데이터

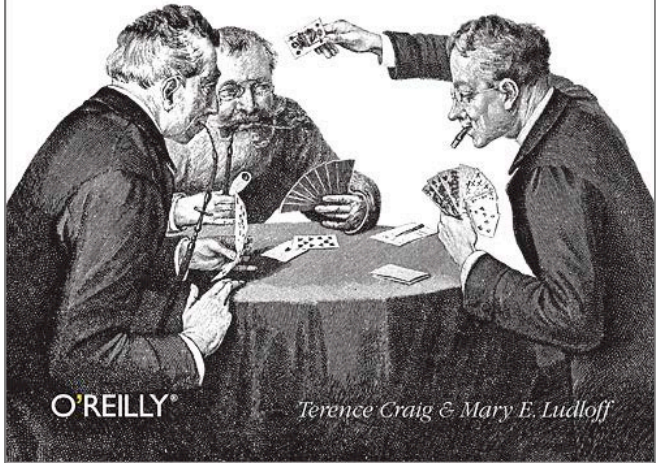
Privacy and Big Data

테렌스 크레이그, 메리 루들로프 지음 / 이춘식, 김정환 옮김

O'REILLY®  한빛미디어
Hanbit Media, Inc.

The Players, Regulators, and Stakeholders

Privacy and Big Data



O'REILLY®

Terence Craig & Mary E. Ludloff

이 도서는 O'REILLY의
Privacy and Big Data의
번역서입니다.

빅데이터 시대의 개인정보보호와 사생활

프라이버시와 빅데이터

빅데이터 시대의 개인정보보호와 사생활 **프라이버시와 빅데이터**

초판발행 2013년 3월 11일

지은이 테렌스 크레이그, 매리 루들로프 / 옮긴이 이춘식, 김정환 / 펴낸이 김태현
펴낸곳 한빛미디어(주) / 주소 서울시 마포구 양화로 7길 83 한빛미디어(주) IT출판부
전화 02-325-5544 / 팩스 02-336-7124
등록 1999년 6월 24일 제10-1779호
ISBN 978-89-6848-603-6 15560 / 정가 9,900원

책임편집 배용석 / 기획 김창수 / 편집 안선화
디자인 표지 여동일, 내지 스튜디오 [맘], 조판 박진희
마케팅 박상용, 박주훈, 정민하

이 책에 대한 의견이나 오타자 및 잘못된 내용에 대한 수정 정보는 한빛미디어(주)의 홈페이지나 아래 이메일로 알려주십시오.
한빛미디어 홈페이지 www.hanb.co.kr / **이메일** ask@hanb.co.kr

Published by HANBIT Media, Inc. Printed in Korea

Copyright © 2013 HANBIT Media, Inc.

Authorized Korean translation of the English edition of *Privacy and Big Data*, ISBN 9781449305000

© 2011 Mary E. Ludloff and Terence Craig. This translation is published and sold by permission of O'Reilly Media, Inc., which owns or controls all rights to publish and sell the same.

이 책의 저작권은 오라일리사와 한빛미디어(주)에 있습니다.

저작권법에 의해 보호를 받는 저작물이므로 무단 복제 및 무단 전재를 금합니다.

지금 하지 않으면 할 수 없는 일이 있습니다.

책으로 펴내고 싶은 아이디어나 원고를 메일(ebookwriter@hanb.co.kr)로 보내주세요.

한빛미디어(주)는 여러분의 소중한 경험과 지식을 기다리고 있습니다. **지금 하지 않으면 할 수 없는 일이 있습니다.**

지은이_ 테렌스 크레이그 Terence Craig

테렌스 크레이그는 빅데이터 분석 서비스와 솔루션을 제공하는 업체인 패턴빌더 PatternBuilders의 CEO이자 최고기술책임자(CTO)다. 패턴빌더는 앞선 분석을 통해 기업이 그들의 사업을 넘어 조직을 이해하고 개선하도록 도와주고 있다. 패턴빌더에서 일하며 테렌스는 공기업과 사기업 모두를 대상으로, 전사적 자원 관리Enterprise Resource Planning, ERP, 전문적인 서비스 자동화Professional Services Automation PSA, 그리고 반도체 공정 제어와 같은 다양한 영역을 넘어 분석 중심의 대규모 사업과 SaaSSoftware as a Service 응용프로그램을 만들고, 시행하고, 판매하는 광범위한 분야에서 활동하였다. 이 책에서 그는 최첨단 기술 회사에서 20년 넘게 운영과 기술 관리를 담당한 경험을 바탕으로, 기업과 조직의 수익성과 효율성을 향상하기 위해 디자인된 세계 수준의 호스팅 분석 플랫폼을 구축하는 데 (운영과 기술 관점에서부터) 무엇이 필요한지에 대한 독특하고 혁신적인 관점을 제시한다. 테렌스는 기술, 착수, 분석, 데이터 보안, 그리고 데이터 프라이버시 윤리와 정책에 대해 알리고 그에 관해 활발히 활동하는 연결가이자, 블로거이며 ‘논평가’기도 하다.

지은이_ **메리 루들로프** Mary E. Ludloff

메리 루들로프는 패턴빌더의 마케팅 부사장으로, 첨단 기술 사업 분야에 20년이 넘는 경험을 지닌 혁신적인 임원이다. 메리는 전통적인 기업과 마케팅 분야로부터 디지털 마케팅의 최신 개발에 이르기까지, 마케팅 분야 전반에 걸쳐 전략적인 프로그램 계획을 어떻게 개발하고 시행하는지 이해하고 있다(그렇다. 그녀의 말보다 훨씬 더 기술적이고 덜 재미있다). 또한 그녀는 브리오 테크놀로지Brio Technology(현재 Hyperion)와 논스톱 솔루션즈NONSTOP solutions(현재 Manhattan Associates)에 근무하면서 데이터 저장과 분석 응용프로그램에 공을 들여, 기업 정보 수집 도구의 사용뿐만 아니라 연쇄적인 생산 및 공급 과정 관리 문제에 대해서도 깊이 이해하고 있다. 활발한 백서 기고자이자 많은 참고 설명서의 편집자로서, 메리는 디지털 시대의 기술과 기업 관점, 데이터 보안, 데이터 프라이버시, 그리고 마케팅으로부터 빅데이터, 분석에 이르는 다양한 쟁점에 대해 출판하는 한편 블로거로도 활동 중이다.

웁긴이_ 이춘식

LG CNS 데이터베이스 관리 팀의 팀장으로 재직 중이다. 데이터베이스를 활용한 개발로부터 설계, 구축과 진단, 컨설팅, 운영까지 다양한 경험을 보유하고 있다. 최고 국가 엔지니어링 자격인 정보관리기술사로서 대학과 기업체에서 강의와 멘토링 활동을 하는 한편, 집필과 기고 활동 역시 폭넓게 수행하고 있다. 최근에는 빅데이터와 전문 데이터베이스인 NoSQL에 관해 번역 및 평역하는 데 관여했고 빅데이터가 실질적인 환경에서 적용되기 위한 방법에 대해, 특히 데이터베이스 관점에서 실증적 적용점을 연구 중이다. 지은 책으로는 『데이터베이스 설계와 구축』(한빛미디어, 2002)외 열 권 정도가 있다.

웁긴이_ 김정환

LG CNS 데이터베이스 관리 팀의 사원으로 재직 중이다. 막 썬낸 빵처럼 시작하는 DBA로써 설레는 마음으로 전문가의 길을 가고 있으며, DBMS의 구조 및 메커니즘을 이해하기 위해 많은 열정과 노력을 기울이고 있다. 최근에는 안정적인 데이터베이스 운영을 위한 트러블슈팅과 최근 업계의 화두로 급부상한 빅데이터의 효율적인 활용 및 이를 위한 다양한 툴을 연구하고 있다.

저자 서문

메리와 내가 책을 쓰는 동안 이 굉장한 생산물을 어떻게 합작했는지에 관한 질문을 여러 번 받았다. 그 다음으로 많았던 질문은 물론 빅데이터 시대의 프라이버시 상태에 대한 우리의 생각을 바꿀 생각이 있는 지에 대해서였다. 우리가 어느 편에 있는지에 상관없이, 책에서 우리는 프라이버시의 양쪽 편을 모두 대변하고자 노력했다. 독자들이 일부 화제에 대한 우리의 생각이 어떤 것인지 알아챌 수 있으리라 확신한다. 이 책은 진행 과정과 프라이버시에 대한 우리의 생각을 일반적인 관점에서 여러분과 공유할 수 있는 기회라고 생각한다.

트렌스 크레이그의 관점

메리와 나는 오랜 친구이자 동업자다. 이번 작업은 우리가 함께하는 두 번째 시도였다. 애초 기술적 설립자이자 CEO인 나와 전통적인 훈련을 받은 마케팅 부사장인 그녀가 잘 해나가지 못하리라는 것은 불 보듯 뻔한 일이었다. 그러나 고맙게도 우리의 경우 서로에 대해 알아가는 것이 두 사람 모두에게 즐겁고 유익한 협업이었다. 어쨌거나 공동으로 이 책을 쓰는 게 얼마나 어려웠는지가 궁금한가? 글을 쓰는 방법과 집필 기한, 스타일을 맞추는 것, 그리고 문법에 대한 서로 다른 해석으로 다루는 것이 정말 힘들었다. 메리는 『Strunk & White』⁰¹의 팬이지만 나는 커밍스^{e.e cummings}⁰²를 신으로 생각한다. 뿐만 아니라 내용 자체에 대해서도 이견이 있었다. 우리가 책에서 다룰 프라이버시는 우리 아버지가 ‘술 먹고 싸우기에 가장 좋은 주제 삼중세트’라고 부를 만큼 논란이 많은 주제다. 예를 들면 종교, 정치, 그리고 다른 남자의 아내가 그런 주제일 것이다(술자리에서라면, 이 세 가지는 심지어 가장 친한 친구 사이에서라도 의자를 휘두르게 할 만한 화제다).

01 글 쓰는 방식에 대한 안내서다. http://en.wikipedia.org/wiki/strunk_%26_white 참조.

02 **역자주** 미국 시인

프라이버시는 우리도 정부도 짜증나게 하는 것 같은데, 사실 보다 광범위한 주제다. 메리와 나 사이의 논쟁은 사실상 내가 걸보기에는 양립할 수 없는 두 가지 입장을 채택한 것을 중심으로 일어나는데, 빅데이터 기술로 인한 프라이버시의 침해가 의미하는 것에 대한 공포와 마크 주커버그⁰³Mark Zuckerberg가 인용한 “프라이버시는 죽었다”에 대한 나의 동의가 그것이다.

어렸을 때 나는 군사 독재정권 국가(나이지리아)에 사는 미국 시민이었다. 1967년 나이지리아 내전⁰³이 발발했을 때, 어머니와 나는 초기에 다른 미국인 여성 및 아이들과 대피했지만 아버지를 포함한 많은 사람들은 그렇지 못했다. 야만스러움으로 유명한 그 내전 기간 동안 그들은 전쟁의 반대 진영에 선 대학생과 근로자를 다락과 지하실에 숨겨주었다. 그리고 나는 그 일을 여전히 자랑스럽게 기억한다.

당시 전쟁으로 인해 2백만 명 이상의 사망자가 발생했는데, 대부분 굶주림 때문이었다. 숨어 있던 사람들이 발견될 경우 그들 모두는 물론 안식처를 제공한 사람들까지 손쓸 틈도 없이 살해되었다. 소련의 감시하에 고통 받던 친구가 많았던 점과 그러한 비극을 본 것은 디지털 프라이버시에 대한 나의 접근 방식을 형성하는 데 도움을 주었다.

나이지리아에서 일어난 일이 여기에서도 일어난다면 어떨까? 2011년 현재, 디지털화된 어떤 나라에서든 저런 피난민과 그들을 숨겨준 용감한 사람들을 간단히 찾을 수 있다. 상대적으로 싼 하드웨어나 우리 회사⁰⁴에서 판매 중인 것과 비슷하며 손쉽게 사용 가능한 상업용 분석 소프트웨어를 사용하여 그들을 찾는 것은, 쉽게

03 http://en.wikipedia.org/wiki/nigerian_civil_war 참조

04 <http://www.patternbuilders.com/>

사용 가능한 일부 데이터 소스를 망가트리는 것에 지나지 않는 요구를 한 것과 같다. 즉, 소셜 미디어, 휴대전화 송수신, 학생과 근로자의 기록을 말한다. 한 번 지원자들이 '찾아질 것' 같으면, 당신은 그들의 거주지를 알아내기 위해 전력과 물의 비정상적인 사용량 증가, 또는 증가한 음식의 양이나 화장실 휴지 구매를 위한 고객 카드 정보의 탐색이 그들의 기록과 관련 있다는 걸 발견할 수 있을 것이다.

이 책을 쓰기 전 디지털 프라이버시에 대한 나의 접근은 가능한 한 인터넷과 떨어진 정보를 유지하는 반면, 가능한 한 부정확성을 유지하지 않도록 계획되었다. 이것이 지나칠뿐더러 피해 망상적이기까지 해서, 가깝고 소중한 나의 많은 지인과 부딪혔다. 내전이 일어난 나라에서 살았던 경험 때문에, 아마도 그들은 이해하지 못하는 것을 나는 빨리 해결할 수 있다. 나는 지인들에게 내가 얼마나 빨리 해결할 수 있는지 그들이 이해할 때까지 답했다. 이 책을 씀으로써 두 가지 정도 흥미로운 방식에 관한 나의 관점이 변했다.

첫 번째는 인정하건대 '패배주의적'이다. 수반되는 모든 불편함을 견디고 기꺼이 완벽하게 자급자족하여 살 수 없다면, 당신은 정부는 고사하고 이웃으로부터도 전통적인 수준의 프라이버시를 유지할 수 없으리라고 나는 믿어왔다. 점점 더 디지털화하는 세계에서 그것은 정말로 불가능하다. 나는 구글Google 맵, 페이스북facebook, 그루폰Groupon, 휴대전화, 그리고 전자 세금 환급을 기꺼이 포기하지 않는다. 그리고 내가 그것을 좋아하든지 안 좋아하든지 관계없이, 공적인 정보와 사적인 정보의 인터넷 추적, DRM, 정보의 융합을 통한 새로운 서비스 제공(매쉬업), 그리고 매우 빠른 속도의 분석 소프트웨어와 하드웨어는 이미 우리 삶의 일부다.

두 번째는 좀 더 '희망적'이다. 최근 위키리크스WikiLeaks가 공개한 미국 정부 기밀의 정확성에 대한 당신의 견해가 어떠한 간에, 그것은 세계적으로 탁월함을 인정 받은 군사력일지라도 유비쿼터스 컴퓨팅의 투명한 유도 효과에 영향을 받는다는 것을 명백히 보여준다. 개인 프라이버시가 침해될 뿐만 아니라 (친구 관계로부터 인류

역사 내내 나타난 부패한 정부, 범죄자, 독재정권까지) 비밀을 유지하기 위한 독재자의 능력도 침식되고 있다. 프라이버시 침해는 비밀스러운 침식의 일부이다. 나의 진실한 희망은 전자에 의해 가능하게 된 잠재된 공포가 후자의 빛에 의해 예방된 공포로 압박받는 것이다. 그리고 나는 예전의 프라이버시 보호 규범으로 회귀하는 것은 꿈에서나 가능하다고 믿는다. 때문에 우리는 내가 옳기를 기원해야만 한다.

그러나 내가 아닌, 책으로부터 기억해야 하는 한 가지는 다음의 문장이다. “인터넷에서 무슨 일이 일어나든지, 그것은 인터넷 밖으로 나와서는 안 된다.”

메리 루들로프의 관점

나는 우리 책이 거의 완료된 시점에 이 글을 쓰고 있다. 지금은 생산 단계고, 트렌스와 나는 대부분의 힘든 집필을 마쳤다. 시간에 관해서라면, 이 책을 쓰기로 서명한 후 실제로는 그리 오래 걸리지 않았다. 처음에 개념을 잡는 것으로부터 시작하여 출판 날짜가 결정되기까지 6개월도 안 걸렸다. 생각과 지력에 대해서는 현재는 뭐, 매우 다른 이야기였다!

참으로 길고, 힘들며, 때로는 격렬한(물론, 가능성 있는 최고의 방법 안에서) 여행이었다. 당신도 알다시피 개인 소유의 작은 회사를 운영한다는 것은 정말 가장 좋은 시간을 의미하고, 당신이 이미 여러 가지 일을 하는 상태에서 책을 쓰는 일을 가장 우선하여 추가할 때, 당신에게는 작은 분열이 생기는 경향이 나타난다. 왜냐하면, 당신은 단지 시간이 없고 심지어 시간이 있을 때도 주로 프라이버시의 해석에 관해 이야기할 것이기 때문에, 당신의 가족과 친구가 아주 조금 당신에게 짜증을 낼 수도 있음을 의미한다. 그래서 나를 이해해주고 우리의 책을 읽고 비평해주어 고맙다는 말을 내 모든 친구와 가족에게 전한다.

트렌스와 함께 이 작업을 시작하면서, 우리 둘은 뭔가 흥미로운 것에 대하여 공동으로 이바지할 수 있다고 생각했다. 우리 사이는, 트렌스와 나는 다른 성별과 다른

역할(마케팅 대 컴퓨터에 매우 능숙한 과학자/기술자/CEO), 그리고 여러 인종을 대표한다. 우리는 매우 다른 출신 배경과 세계관을 가지고 있다. 특히 프라이버시에 대해 그렇다. 우리 둘이 그 주제에 대해 아무리 많이 대화하고 블로그에 기록한다 해도, 우리 각자는 다른 사람에게 “당신은 핵심을 놓치고 있어요”라고 말하는 것으로 알려지는 것이 안전할 것이다. 우리는 함께 판단했고, 그 결과 프라이버시에 대한 관점을 꽤 많이 보여줄 수 있었다. 서로 다른 우리의 관점은 몇몇 흥미로운 논의를 만들 어낼 수도 있을 것이다. 그리고 실제로 그렇게 되었다.

내가 계산하지 못했던 것은 책을 집필하는 것이 프라이버시에 대한 나의 관점에 어떤 영향을 끼치는가 하는 점이었다. 지금 당신이 우리 블로그(<http://blog.patternbuilders.com>)를 방문한다면, 그에 대한 글을 자주 올리기 때문에, 당신은 아마 프라이버시 논쟁에 대한 나의 견해를 꽤 친근하게 느낄 것이다. 나의 관점에 익숙하지 않다면, 여기 짧은 버전을 참고하기 바란다.

- 미국은 좀 더 포괄적인 프라이버시 보호 법안을 제정할 필요가 있으며 강력한 실행력을 갖추는 것도 필요하다.
- 개인 정보를 수집하고 빌리는/파는 자는 누구든지 항상 사용자에게 대해 알리고 모든 데이터 사용은 반드시 사전에 동의를 받아야 한다.
- 프라이버시 정책은 표준화되어야 한다. 특정한 제삼자가 사용하는 것을 포함하여, 표준화되지 않은 프라이버시와 관련된 어떠한 것도 반드시 설명되어야 하며 사전에 제공되어야 한다.

꽤 간단하지 않은가? 그런 프라이버시를 제외하는 것은 단순한 주제가 아니다. 프라이버시는 복잡하고 미묘하며 그것에 대한 상당히 많은 일면이 있다. 우리가 온라인에서 다양한 활동을 할 수 있도록 기술은 새로우면서도 다양한 방법을 계속 제공하고 있으며, 데이터는 경계가 없으나 프라이버시 규제는 있다는 사실 등, 이쯤이면 우리가 손을 올리고 “항복!”이라고 말하기에 충분하다.

트렌스와 집필을 시작할 때, 나는 책의 결말을 꽤 확신했다. 온라인상에는 수많은 우리의 개인 정보가 떠돌고 있다. 우리는 매우 암울한 디지털 시대에서 그것이 어떻게 사용되는지, 개인의 프라이버시 유지에 대한 어떠한 관점이 만들어지고 있는지 거의 모른다. 그러나 나는 그 관점이 장밋빛이 아닐지라도, 우리 각자는 다음에 해야 하는 것을 제어할 수 있다는 것을 발견했다.

우리의 개인 정보가 온라인상에 떠도는 것은 기정사실이지만(믿어지지 않는다면 '스포키오⁰⁵spokeo'에서 당신 이름을 검색해보라) 우리는 스마트폰, 아이패드, 노트북, 또는 스스로가 좋아하는 기록 장치를 이용할 때마다 얼마나 많은 개인 정보가 추가되는지 파악하고, 계속 통제한다. 이제, 당신이 온라인 상태일 때 프라이버시의 수준이 어느 정도인지에 대해 생각해보고 이 시간부터 무엇을 할 것인지 결정해보라(반약 당신이 현재 상태에 만족한다면, 그대로 유지하라).

- 페이스북 사용하지 않기. 나는 한 번도 계정을 가진 적이 없고 앞으로도 이용하지 않기로 했다. 혹시 페이스북이 프라이버시 분야에서 (조심스럽게 말하면) '뛰어나지' 않기 때문에 내가 그렇게 결정했다고 생각하는가? 그렇다면 당신은 틀렸다. 나는 내 삶을 오프라인으로 유지하기로 오래 전에 결정했고(내 직업적인 것은 어디에나 있다), 현재 그렇게 유지하고 있다.
- 지독하게 프라이버시를 침해하는 기업과는 거래하지 않기. 그 기업들이 스스로의 행동을 청산하고 다시금 바른 생활을 하고 있음을 나에게 입증할 때까지 거래하지 않는다.
- 프라이버시 보호 인증을 획득했거나, 상품에 프라이버시 보호를 내재하고 있거나, 프라이버시 문제에 대해 빠르게 답변하는 (그리고 개선하는) 프라이버시 방침을 따르는 기업과 거래하기(누구나 실수할 수 있기 때문에).

05 역자주_신상정보를 모집하는 사이트. <http://www.spokeo.com/>

- 온라인 사이트에 개인적인 사진이나 비디오 같은 ‘개인적인’ 것 올리지 않기. “이봐, 이걸 모두를 위한 것이 아니고 내가 살아가는 방침이야(그렇다, 가족과 친구들 때문에 나는 힘든 시간을 견뎌야 했지만, 결국 그들은 그들의 페이스북 페이지나 다른 인터넷상에 나를 포함하지 않는 친절함을 베풀었다).”
- 프라이버시 보호 활동가 되기. 무슨 일이 일어나는지 관심이 없다 하지만, 나는 트위터나 블로그 등에서 그런 점에 관한 대화를 하고 있다. 현재 우리가 사는 세계의 위대한 점은 소셜 미디어를 통해 우리가 모든 것을 들을 수 있다는 점이다.

들어보라. 도구를 사용하는 일로부터 다가오는 온라인 시대를 단순히 맞이하지 않는 것까지, 프라이버시의 침해를 완화하기 위해 우리가 할 수 있는 일들은 존재한다. 우리는 우리가 프라이버시 인증(Truste⁰⁶와 같은)을 찾는 것을 믿는 사람이나 프라이버시 행동강령에 헌신하는 사람, 또는 그들의 제품에 프라이버시를 넣는 사람들(Privacy by Design⁰⁷)에게 우리의 사업을 넘길 수도 있다. 기업체가 나쁘게 행동할 때, 꼭 법적인 방법이 아니라도 우리 스스로 그들을 처벌할 수 있다. 더는 그 사이트를 이용하지 않거나 우리의 회원권을 폐기하는 것이다. 패배를 인정하는 대신(디지털 세상에서 프라이버시와 같은 것이 없다면) 나는 예전보다 더 열정적이다.

이 논쟁에서 상대가 알아듣게끔 우리가 이야기해줄 만한 시간은 여전히 있고, 의미 있는 변화를 위한 시간도 있다. 그러나 디지털 시대에서 프라이버시가 의미하는 바를 정확하게 알아가기 시작하는 것은 우리, 나, 당신 그리고 모두에게 달린 문제다. 그런 다음 어떻게 할지에 대해서는 리얼리티프로그램인 ‘프로젝트 런웨이 Project Runway’의 진행자 팀 건 Tim Gunn의 말을 인용한다. “작동하게 만드세요.”

06 <http://www.truste.com/>

07 <http://privacybydesign.ca/>

트렌스와 나는 오랜 대화를 통해 프라이버시에 대한 견해를 나누었다. 책의 마지막 장에서, 그와 내가 공유했던 것을 당신과도 나눌 것이다. 이것이 나의 꿈이다(시작하려는 아이디어를 찾는 사람들은 주목하라). 만약 마이크로소프트와 다트머스대학이 인터넷에서 아동의 성적 착취 이미지를 제거하는 데 도움이 되는 PhotoDNA⁰⁸를 개발할 수 있다면(이것은 놀라운 이야기다. 이 프로젝트에 대해 들어본 적이 없다면, 링크를 따라가서 그에 관한 정보를 읽기 바란다), 5년 뒤쯤에는 우리의 정보가 저 지점(또한 어떻게 사용되는지)에서 인터넷 전반으로 추적할 개인 정보 DNA를 찾아낼지 누가 알겠는가?

그렇다면 우리는 스스로의 개인 정보를 나누어줄 때 정확히 무슨 일이 일어날지 볼 수 있거나, 나의 시나리오상에서 프라이버시 보호에 관한 '에퀴팩스 Equifax'⁰⁹ 버전이 되기 위해, 몇몇 기업에 월 20달러를 지급 할 수 있을 것이다(내 프라이버시가 침해당하면 나에게 경고하는 식으로).

이와 같은 일이 절대로 일어나지 않을 것이라 생각하는가? 그렇다면 현재 당신의 삶에 활력을 주는 모든 장치에 대해 생각해 보라. 그 중 많은 부분이 5년 전에 존재하지 않았고, 대부분은 10년 전에는 존재하지 않았을 것이다. 5년 뒤 프라이버시 환경이 어떻게 변할지 누가 알겠는가? 한 가지 내가 확신하는 것이 있다. "나는 다음 번에 무슨 일이 일어날지 계속 주시할 것이다!"

08 <http://www.microsoft.com/presspass/presskits/photodna/>

09 **역주** 미국 신용평가, 조사 전문업체다. 신원 확인 관리의 리더로 불린다. <http://www.equifax.com/>

역자 서문

길을 가다가 가끔 멈추어 ‘내가 왜 이 길을 가고 있지?’라는 의문을 마음속에 떠올릴 때가 있다. 정신없이 무언가에 몰입하다가 갑자기 ‘그런데 내가 왜 이 일을 하고 있지?’라는 생각을 할 때도 있다.

빅데이터를 정신없이 외치는 최근 세대에는 짬, 생각, 물러섬Retreat, 점검 등의 단어가 필요할 것인데, 이 책은 이런 유의 단어에 어울릴만한 책이다. 정신없이 쏟아지는 ‘빅데이터’ 관련 정보의 홍수 속에서 간과할 수 없는 이슈는 바로 ‘개인 정보’와 ‘프라이버시’다. 어차피 데이터를 활용한 다양한 시도는 계속 증가할 것이고, 데이터를 통한 비즈니스 가치 증대는 거대한 물결이 되어 이미 진행 중이다. 항상 그렇지만 앞만 보고 정신없이 달려가다 보면 가끔 잘못된 길로 들어설 수도 있고, 또는 검증하고 가야 할 사항들을 간과해버려서, 결과적으로 치명적인 문제가 기업에 나타날 수 있다.

이 책 한 권으로 데이터와 관련된 프라이버시 문제를 모두 해결할 수는 없다. 그러나 적어도, 이 책을 통해 빅데이터를 적용할 때 고려해야 할 프라이버시는 무엇이 있고 어떤 관점으로 적용해야 하는지 정리해서 질문할 수는 있을 것이다. 앞만 보고 달려가는 빅데이터와 관련된 모든 사람에게, 한 가지 꼭 주의해서 알고 가야 할 사항에 대해 ‘이것은 꼭 챙겨주세요!’라고 요청하는 책 씬으로 생각해주면 만족한다.

‘지독하게 프라이버시를 침해하는 기업과는 거래하지 않기-그들이 자신의 행동을 청산하고 다시금 바른 생활을 하고 있다는 걸 나에게 입증할 때까지 거래하지 않는다.’ 본문에 있는 내용이다. 앞으로 우리가 비즈니스 관계를 맺는 다양한 영역에서 어떤 기업이 프라이버시를 소홀히 한다면 누군가는 이런 생각을 할지도 모를 일이

다.

빅데이터와 관련한 논의를 계속하고 싶다면 IT분야 사람들이 많이 모여 있는 네이버 카페 '아이리포(<http://cafe.naver.com/itlf>)'를 찾아보기 바란다.

이 책이 나올 수 있도록 지원해준 새로운 생각과 열정을 가진 LG CNS의 DB 관리 팀 사람들에게 감사의 말을 전한다. 또한, 항상 혁신을 이끌어내는 한빛미디어 스마트미디어팀에게 감사드린다.

2013년 2월
이춘식

한빛 eBook 리얼타임

한빛 eBook 리얼타임은 IT 개발자를 위한 eBook 입니다.

요즘 IT 업계에는 하루가 멀다 하고 수많은 기술이 나타나고 사라져 갑니다. 인터넷을 아무리 뒤져도 조금이나마 정리된 정보를 찾는 것도 쉽지 않습니다. 또한 잘 정리되어 책으로 나오기까지는 오랜 시간이 걸립니다. 어떻게 하면 조금이라도 더 유용한 정보를 빠르게 얻을 수 있을까요? 어떻게 하면 남보다 조금 더 빨리 경험하고 습득한 지식을 공유하고 발전시켜 나갈 수 있을까요? 세상에는 수많은 종이책이 있습니다. 그리고 그 종이책을 그대로 옮긴 전자책도 많습니다. 전자책에는 전자책에 적합한 콘텐츠와 전자책의 특성을 살린 형식이 있다고 생각합니다.

한빛이 지금 생각하고 추구하는, 개발자를 위한 리얼타임 전자책은 이렇습니다.

1. eBook Only - 빠르게 변화하는 IT 기술에 대해 핵심적인 정보를 신속하게 제공합니다.

500페이지 가까운 분량의 잘 정리된 도서(종이책)가 아니라, 핵심적인 내용을 빠르게 전달하기 위해 조금은 거칠지만 100페이지 내외의 전자책 전용으로 개발한 서비스입니다. 독자에게는 새로운 정보를 빨리 얻을 수 있는 기회가 되고, 자신이 먼저 경험한 지식과 정보를 책으로 펴내고 싶지만 너무 바빠서 엄두를 못 내시는 선배, 전문가, 고수분에게는 보다 쉽게 집필하실 기회가 되리라 생각합니다. 또한 새로운 정보와 지식을 빠르게 전달하기 위해 O'Reilly의 전자책 번역 서비스도 준비 중이며, 조만간 선보일 예정입니다.

2. 무료로 업데이트되는, 전자책 전용 서비스입니다.

종이책으로는 기술의 변화 속도를 따라잡기가 쉽지 않습니다. 책이 일정한 분량 이상으로 집필되고 정리되어 나오는 동안 기술은 이미 변해 있습니다. 전자책으로 출간된 이후에도 버전 업을 통해 중요한 기술적 변화가 있거나, 저자(역자)와 독자가 소통하면서 보완되고 발전된 노하우가 정리되면 구매하신 분께 무료로 업데이트해 드립니다.

3. 독자의 편의를 위하여, DRM-Free로 제공합니다.

구매한 전자책을 다양한 IT기기에서 자유롭게 활용하실 수 있도록 DRM-Free PDF 포맷으로 제공합니다. 이는 독자 여러분과 한빛이 생각하고 추구하는 전자책을 만들어 나가기 위해, 독자 여러분이 언제 어디서 어떤 기기를 사용하시더라도 편리하게 전자책을 보실 수 있도록 하기 위함입니다.

4. 전자책 환경을 고려한 최적의 형태와 디자인에 담고자 노력했습니다.

종이책을 그대로 옮겨 놓아 가독성이 떨어지고 읽기 힘든 전자책이 아니라, 전자책의 환경에 가능한 최적화하여 쾌적한 경험을 드리고자 합니다. 링크 등의 기능을 적극적으로 이용할 수 있음은 물론이고 글자 크기나 행간, 여백 등을 전자책에 가장 최적화된 형태로 새롭게 디자인하였습니다.

앞으로도 독자 여러분의 충고에 귀 기울이며 지속해서 발전시켜 나가도록 하겠습니다.

지금 보시는 전자책에 소유권한을 표시한 문구가 없거나 타인의 소유권한을 표시한 문구가 있다면 위법하게 사용하고 계실 가능성이 높습니다. 이 경우 저작권법에 의해 불이익을 받으실 수 있습니다.

다양한 기기에 사용할 수 있습니다. 또한 한빛미디어 사이트에서 구입하신 후에는 횡수에 관계없이 다운받으실 수 있습니다.

한빛미디어 전자책은 인쇄, 검색, 복사하여 붙이기가 가능합니다.

전자책은 오타자 교정이나 내용의 수정보완이 이뤄지면 업데이트 관련 공지를 이메일로 알려드리며, 구매하신 전자책의 수정본은 무료로 내려받으실 수 있습니다.

이런 특별한 권한은 한빛미디어 사이트에서 구입하신 독자에게만 제공되며, 다른 사람에게 양도나 이전되지 않습니다.

차례

01	퍼펙트 스톰	1
	1.1 거울을 통해	4
	1.2 빅데이터 시대에 오신 것을 환영합니다!	6
	1.3 퍼즐 조각에서부터 완성된 그림까지: 미래는 바로 지금이다.	8
	1.4 아주 위험한 적과 같은 광고	9
	1.5 세계를 장악한 빅브라더와 빅데이터	12
	1.6 프라이버시 대 보안과 안전, 그 갈림길에서?	15
	1.7 참고 문헌	18
02	디지털 시대의 프라이버시 권리	21
	2.1 디지털 시대에 프라이버시의 의미는 무엇인가	22
	2.2 미국에서의 프라이버시: 홀로 있을 권리	26
	2.3 유럽에서의 프라이버시: 명예와 존엄성	28
	2.4 우리는 늘 어떤 프리즘을 통해 프라이버시를 본다	32
	2.5 국경을 초월하는 프라이버시	33
	2.6 가치관의 충돌	34
	2.7 서로 연결된 프라이버시: ‘나’와 집단으로서의 ‘우리’	36
	2.8 참고 문헌	37
03	규제	40
	3.1 ‘디지털’ 프라이버시 규제의 짧은 역사	43
	3.2 프라이버시 규제 모델: 상호보완적인가, 모순적인가	44

3.3	미국의 규제 모델: 상향식 접근	46
3.3.1	미연방통상위원회(FTC)	47
3.3.2	미연방통신위원회(FCC)	49
3.3.3	상무부(Commerce)	50
3.3.4	에너지부(DOE)	50
3.3.5	보건사회복지부(HHS)	51
3.3.6	소비자금융보호국(CFPB)	51
3.3.7	미국 모델에 대한 마지막 의견	52
3.4	유럽연합 규제의 모델: 하향식 접근	52
3.4.1	안전피난처 효과	55
3.4.2	유럽연합 모델에 대한 마지막 의견	56
3.5	다른 국가의 프라이버시 보호법에 대한 간략한 개론	57
3.6	프라이버시 대 보안과 안전	58
3.7	데이터는 죽지 않는다	60
3.8	정부기관은 프라이버시 권리에 영향을 받는가	60
3.9	참고 문헌	61

이해 당사자들 68

4.1	이해 당사자는 누구인가	72
4.2	온라인 광고의 간략한 역사	73
4.3	지적재산권, 신뢰 컴퓨팅, 그리고 프라이버시	77
4.4	거대 기업과 프라이버시	84
4.5	빅데이터의 출현과 프라이버시	89
4.6	지니를 램프 밖으로 소환하다	93

4.7	프라이버시 수호를 위한 단체의 노력	98
4.8	떠오르는 프라이버시 경제	100
4.9	프라이버시 수호를 위한 개인의 노력	104
4.10	참고 문헌	105

그 모두에 대해 이해하기

5.1	문제의 핵심: 프라이버시를 상품으로 볼 것인가, 권리로 볼 것인가	114
5.2	우리는 모두 연결되어 있다	115
5.3	안전과 보안을 위해 무엇을 흔쾌히 포기할 수 있는가	118
5.4	데이터에 관한 진실: 한 번 유출되면 통제하기 어렵다	121
5.5	다시 처음으로	122
5.6	참고 문헌	123

최근 20년간 첨단기술 사업 분야에서 일한 사람이라면, 지속적으로 진화하는 데이터 프라이버시 논쟁에 대해 조감도를 보듯 전체를 바라볼 수 있는 안목을 지녔을 것이다. 개인의 데이터가 어떻게 수집되고 활용되느냐에 대해 무관심한 시각도 있고, 개인의 정보가 하이재킹(hijacking)당하는 방법이 거의 무한대로 존재할 것이라 보는 보다 냉소적이고 편집증적인 관점도 존재한다. 이 중 당신이 어떤 입장에 서든간에, 그 위험성만큼은 사상 최대라고 말할 수 있다.

퍼펙트 스톰이 형성되고 있다. 이 폭풍우의 연료는 우리가 이야기하고 의사소통하는 방법에 붙어 닥친 '혁명적인 변화'다. 인터넷이 우리의 삶에 이다지도 전반적으로 영향을 미칠 것이라고, 20년 전이라면 누군들 예측할 수 있었을까? 잠자는 시간을 빼고 우리는 거의 매일을, 24시간 내내 웹 사이트에 접속하고 있다. PC나 스마트폰, 아이패드 등으로 이메일을 확인하고, 블로그의 흥미로운 글을 읽고, 식당을 찾고, 일자리를 구하고, 친구들의 페이스북 담벼락을 보고, 책을 사고, 돈을 이체한다. 또, 어디로 가야 할지 방향을 찾고, 트위터나 포스퀘어(Foursquare)에서 자신의 현재 위치를 남기고, 독재에 대항하는 반대 시위를 조직하기도 한다. 세계 어느 곳에 있든지 상관없이 말이다. 우리는 디지털 시대에 살고 있다.

디지털 기술로 말미암아 이전에는 불가능했던 많은 일들이 가능해지는 새로운 세계 질서가 창조되고 발전해왔다. 비록 개인용 제트팩(Jet Pack⁰²)이나 하늘을 나는 자

01 **역자주** 개별적으로는 위력이 크지 않은 태풍 등이 다른 자연현상과 동시에 발생하여 엄청난 파괴력을 내는 현상을 일컫는 용어로, 프리랜서 기자이자 작가인 세바스찬 용거가 1991년 미국 동부 해안에서 벌어진 실화를 바탕으로 쓴 베스트셀러 『퍼펙트 스톰』에서 처음 썼다. 용거는 당시 허리케인 그레이스와 다른 두 개의 기상전선이 충돌해 유례없는 대형 폭풍이 만들어진 걸 보고 '퍼펙트 스톰'이라 이름 지었다.

02 **역자주** 등에 매는 개인용 분사 추진기로 우주 유명 등에 사용된다.

동차를 가질 수 있는 정도는 아니지만, 우리는 지금 영상통화가 되는 전화기를 사용하고 무인정찰기로 전투를 한다. 이런 변화에 따라 개인 프라이버시 영역에도 변화가 생겼다. 영국 작가 조지 오웰(1903~1950, 영국, 『동물농장』 등)이 디스토피아적 세계를 그린 소설 『1984』에서 예견했던 세계 정도는 아직 아니겠지만(소설 속에서는 개인의 사생활을 보호하는 프라이버시 권리가 아예 없고, 정부는 국민을 통제하기 위해 감시하며, 잘못된 정보를 이용하는 것으로 그려진다), 우리 정부 역시 확실히 우리의 개인 정보를 그들의 이익을 위해 이용하고 있으며 조지 오웰이 예상한 것 이상으로 우리에게 관한 많은 것을 알고 있다.

어떤 것은 좋은 쪽으로 또 어떤 것은 나쁜 쪽으로, 세상은 계속해서 변화해왔다. 오늘날 우리는 인간 역사상 그 어떤 시대보다 우리 자신에 대한 정보를 내어놓고 우리의 정보가 수집되고 집약되는 시대에 살고 있다. 그 정보의 대부분은 간단한 편의를 위해 우리 스스로가 제공하는 것이다. 그것도 자유롭게 사용할 수 있도록 ‘무료’로 제공한다. 정보의 일부는 부정한 방법으로 수집되기도 한다. 또한, 2009년에 미국 법무부가 GPS 기능 휴대폰 가입자 8백만 명의 위치 정보를 스프린트(Sprint, 미국 이동통신 회사)로부터 받았던 예⁰³에서 보듯, 정부가 어떤 조치를 위해 정보를 이용하기도 한다.

이제는 오프라인 생활조차도 온라인이라 할 수 있다. 우리는 전자상거래나 메신저를 사용하기 위해, 수많은 친구나 비즈니스 관계의 사람들과 인연을 유지하기 위해, 동창과의 교류를 위해, 관심을 두는 분야의 커뮤니티와 관계를 형성하기 위해, 심지어는 사랑을 얻기 위해 우리의 개인 정보를 온라인에서 교환한다. 이러한 데이터는 집약되어 다양한 기준으로 분류된다. 예를 들면 나이, 성별, 소득 수준, 거주하는 지역, 선호, 방문하는 사이트 등이다. 이 정보는 우리의 행동 관점에 따라 그룹화되고 우리에게 무엇인가를 판매하려는 광고주에게 임대되거나 판매되고 있다.

03 <http://www.wired.com/threatlevel/2009/12/gps-data/>

프라이버시에 대한 논쟁의 대부분은 ‘행동 타겟팅(Behavioral Targeting)’⁰⁴에 집중되어 있다. U.C. 버클리대학과 펜실베이니아대학의 최근 연구⁰⁵에 의하면 설문조사에 응한 사람 중 66퍼센트가 마케터가 그들의 관심사에 따라 광고를 개인화하는 것을 원하지 않는 것으로 나타났다. 설문에 참여한 사람들에게 그들의 행동이 추적된 방법을 들려주자 반대 의견의 비율이 무려 73퍼센트에서 86퍼센트까지 치솟았다. 최근의 오페라소프트웨어 사의 설문조사⁰⁶에 의하면, 미국인들은 테러리스트의 공격이나 개인 파산, 또는 가정 침입보다 온라인상에서의 프라이버시 침해를 더 두려워하는 것으로 조사되었다.

이제 표적 광고(Targeted advertising)⁰⁷는 전혀 새로운 개념이 아니다. 오늘날에는 디지털을 통해 모든 것을 추적하고 그것들을 분류함으로써, 우리가 무엇을 구매할지에 대해 학습된 추측을 하기가 훨씬 쉬어졌다. 그러나 예전보다 더 거슬린다고 해서 광고가 진정 두려워해야 할 대상일까? 우리는 훨씬 다양한 행동에 대해 학습된 추측이 같은 프로세스로 적용될 수 있다는 사실을 생각해보아야 한다. 보안기관은 잠재적 테러리스트의 프로필을 식별하기 위해, 국세청은 허위 세금 환급을 식별하기 위해, 사법당국은 잠재적 범죄 행위를 감시하기 위해, 신용카드사와 대부분체는 신용 위험도를 판단하기 위해 학습된 추측을 이용할 수 있다. 데이터는 그 자체로는 유순하지만, 그것이 이용되는 방법에 따라서 해가 전혀 없는 것으로부터 아주 해로운 경우, 심지어는 완전히 최악이라 불리는 것으로까지 다양한 결과를 낳을 수 있다.

데이터 프라이버시를 이야기할 때, 광고가 어떻게 이루어지는지는 우리에게 논쟁거리가 아니다. 논쟁의 핵심은 상업적이고 정치적 관점에서 우리의 개인 정보를 모

04 **역자주** 광고의 효과를 제고하기 위하여 매체나 광고주가 사용하는 기법으로, 이용자의 방문 페이지, 검색 이력 등과 같이 브라우징 행위에 관해 수집된 정보를 사용한다.

05 http://www.google.com/hostednews/afp/article/ALeqM5jm38cd0yGVDPXRO1_dohxoq__-rw

06 <http://www.opera.com/press/releases/2011/01/28/>

07 **역자주** 이용자의 행태 정보를 활용한 맞춤형 광고로, 이용자가 웹 사이트에 방문하여 수행하는 행태 정보가 일정 기간 동안 수집·저장·추적·분석되고 이를 기반으로 이루어진다.

오고 이용한다는 점이다. 제품 및 서비스 이용의 편의를 위해 우리의 정보를 제공한 일이 결과적으로, 국가와 자치단체, 혹은 치안기관 등이 다양한 명목으로 우리를 감시할 수 있게끔 문을 열어준 것이 되었다. 지금 이러한 결과에 도달한 것은 기술적인 혁신과 기업이 정신에 의한 결과다. 이러한 상황으로부터 우리가 어디로 가느냐 하는 것은 우리 자신에게 달려 있다.

1.1 거울을 통해

모든 것은 1969년 알파넷^{ARPANET}⁰⁸이 설립되면서 시작되었다. 알파넷은 국방부가 군사시설 간 정보의 흐름을 보호하기 위해 지리적으로 분산한 컴퓨터 네트워크로, 네트워크들의 네트워크⁰⁹인 인터넷의 시초다. 지금은 민간, 공공, 정부, 기업 그리고 학술 네트워크에 이르기까지 방대한 양의 디지털 정보를 연결하는 중심체가 되었다.

그러던 중 우리의 일과 개인 생활의 중심에 인터넷이 들어서도록 한 몇 가지 변곡점이 나타난다.

- **인터넷이 일상적인 용어가 되었다** 팀 버너스리^{Tim Berners-Lee}가 웹^{World Wide Web, WWW}에 대한 초기 사상을 만든 1990년 이후 1993에는 DEC^{Digital Equipment Corporation}가 최초의 상업적인 웹 사이트를 공식적으로 출범하였다. 1990년대 중반에는 웹 브라우저가 소개되었고, 이때 이미 근로자 세 명 중 두 명, 세 가 정 중 한 가정은 인터넷에 접근할 정도로 PC를 통한 접속이 급격히 증가할 것이라 예견되었다.
- **온라인을 통해 쇼핑하게 되었다** 1995년에 들어서는 이베이^{ebay}와 아마존^{Amazon}이 설립되어, 온라인에서 소비자와 직접 거래하는 새로운 비즈니스 모델을 만

08 **역자주** ‘Advanced Research Projects Agency Network’의 약자로, 미국 국방부의 고등연구기획청이 자금을 대서 개발한 분산 관리 방식의 통신 네트워크다. 현재 인터넷의 모체가 되었으며, 이에 관한 자세한 내용은 다음의 웹 사이트를 참조하라. <http://en.wikipedia.org/wiki/arpnet>

09 **역자주** 네트워크 간의 상호 연결성 부여

들었다. 이는 이후 자포스Zappos¹⁰나 넷플릭스Netflix¹¹ 같은 새로운 온라인 전용 기업이 출현하는 토대가 되었을 뿐 아니라, 전통적인 굴뚝 산업들이 그들의 비즈니스를 온라인 세상에 맞게 변형하도록 하였다.

- **검색이 대세가 되고 강력하고 새로운 광고 모델임이 검증되었다** 야후와 라이코스 의 뒤를 이어 1998년에는 우수한 광고 타깃팅 기법과 뛰어난 검색 알고리즘 을 무기로 한 구글이 설립되었다. 구글은 사람이 정보를 검색하는 방법을 바꾸 었을 뿐만 아니라, 콘텐츠 기반의 유료 검색 광고 모델을 완성했다.¹² 그 결과 2010년 4분기 구글의 매출은 84.4억 달러에 달했다. 구글은 또한 개인의 행 동에 대해 역사상 가장 많이 집적된 데이터를 만들어내었다고 할 수 있다.
- **소셜 미디어 사이트들이 출현했다** 2003년에는 소셜 네트워크의 개척자들이 나 타났다. 프렌스터(지금은 소셜 게임 사이트Friendster)와 마이스페이스myspace는 페 이스북에게 추월당하기 전까지 가장 인기 있는 소셜 네트워크였다. 2004년 에 소셜 미디어라는 용어가 신조어로 등장하였고(크리스 샤플리Chris Sharpley가 처 음 사용했다) 페이스북이 설립되었다. 2005년 유튜브가 온라인에 등장한 것에 이어 2006년에는 트위터가 등장했다. 이들을 비롯해 많은 사이트들이 개인의 행동과 그들의 위치 정보, 그리고 사람들의 관계에 대해 방대한 양의 디지털 데이터를 만들어내고 있다.
- **개인 단말기가 증가했다** 1996년에 인터넷 접속이 가능한 최초의 모바일 폰인 ‘노키아 9000 커뮤니케이터’가 등장했다. 2001년 블랙베리가 최초로 이메일 이 가능한 휴대폰 시스템을 출시한 데 이어, 2007년에는 애플이 아이폰을 소 개했다. 2008년에 들어서는 인터넷이 가능한 휴대폰이 PC보다도 많아졌다. 2010년에는 아이패드를 필두로 한 태블릿 기기들이 시장을 폭풍처럼 휩쓸었

10 <http://www.zappos.com/>

11 <http://www.netflix.com/>

12 <http://digitalenterprise.org/models/models.html>

고, 그로 인해 더 많은 데이터가 발생하게 되었다. 이제는 사용자의 위치 정보가 기기 자체의 핵심적인 구성 요소가 되었다. 누가 어디에 있는지를 그 자신이 이야기하지 않아도 언제든 알 수 있게 된 것이다.

- **커뮤니케이션이 실시간으로 이루어진다** 1996년에 소개된 AOL의 메신저 IM(Instant messenger)은 실시간으로 메시지를 주고받는 서비스다. 마이크로소프트의 MSN 메신저와 스카이프는 비즈니스 이해관계자나 많은 개인과 연락할 수 있는 수단이다. 단문서비스인 SMS 프로토콜이 1984년에 개발되면서부터는 무선기기로 문자메시지를 보낼 수 있게 되었다. 이제 10대나 젊은 성인은 SMS를 이용한 커뮤니케이션을 더 선호하게 되었다. 2014년 말에 이르면, IM을 이용하는 계정의 수가 35억 개를 넘어설 것으로 추정되고 있다.¹³ 소셜 미디어 사이트와 비슷하게 IM도 엄청난 양의 정보들을 쏟아내고 있는데, 그 정보의 양은 가령 사용자 개인에 대한 정보뿐만 아니라 다른 사람들이나 조직과의 관계의 깊이나 수준을 소셜 그래프로 나타낼 수 있을 정도다.

이제 우리는 항상 네트워크에 연결되는 디지털 세상을 살고 있다. 온라인으로 일하고, 온라인으로 사회생활을 하고, 뉴스를 보고, 좋아하는 프로그램도 본다. 온라인으로 세금을 내고 은행 일을 보며, 심지어 도박도 하고, 성적인 관심을 온라인을 통해 추구하기도 한다. 우리가 디지털 세상에 흔적을 남긴 모든 것, 그것에 이름을 붙인 것이 바로 ‘빅데이터’다.

1.2 빅데이터 시대에 오신 것을 환영합니다

지난 몇 년간 계속 잠들어 있었던 것이 아니라면 디지털 세상에서 생성되는 데이터의 양에 대해 알고 있을 것이다. “데이터에 익사하다”, “데이터의 광대한 대양”, “기하급수적인 데이터의 증가” 등은 빅데이터의 크기를 묘사하기 위해 언급된 구절

13 <http://www.radicati.com/?p=5290>

들이다. 왜일까? 이해하기에는 데이터의 크기가 너무 크기 때문에 그렇다. 아래는 IDC Research의 내용이다.¹⁴

- 2009년에 들어 디지털 세계는 62퍼센트, 숫자로는 거의 80만 페타바이트 (petabytes, 10^{15} , 1,000,000,000,000,000바이트) 가량 증가했다(DVD를 쌓는다면 지구부터 달까지 왕복할 수 있는 양으로, 1페타바이트는 약 100만 기가바이트다).
- 2010년에는 1,200,000페타바이트¹⁵(아직 최종 집계되지 않음)로 증가할 것으로 예상된다.
- 2020년에 이르면 2009년에 비해 44배 더 증가할 것으로 예상된다(DVD를 쌓는다면 지구에서 화성까지 거리의 절반 정도 되는 만큼 쌓일 것이다).

그러나 빅데이터를 다룰 때 단순히 그 크기에 관해서만 이야기하지는 않는다. 사용 가능한 데이터 소스의 순수한 수, 다양한 형태, 그리고 대부분은 사용자에게 의해 생성된다는 사실이 중요하다. 디지털 세계의 70퍼센트는 실제로 우리 모두가 사용하는 이메일, 페이스북, 트위터, 링크드인LinkedIn, 플리커Flickr, 유튜브 등을 통해 생성된 것이고, 이런 항목들은 계속 늘어나고 있다.¹⁶ 다음을 보자.

- 구글 인덱스에 들어 있는 1조 개의 유일한 URL들과 매일 구글에서 행해지는 20억 건의 검색
- 유튜브에서 감상할 수 있는 7천만 개의 동영상(하루를 기준으로 1억 번 재생된다)
- 1억 3천3백만 개의 블로그
- 290억 개 이상의 트위터 메시지(매일 3백만 개씩 증가)
- 5억 명 이상의 페이스북 사용자들과 그들이 한 달간 페이스북을 이용하는 7천억 분 이상의 시간

14 <http://www.emc.com/collateral/demos/microsites/idc-digital-universe/iview.html>

15 <http://en.wikipedia.org/wiki/petabytes>

16 <http://www.emc.com/collateral/demos/microsites/idc-digital-universe/iview.html>

그 외 미국 연방정부, 주정부 그리고 지방정부의 기관, 학교와 연구 기관, 특정 지역과 관련된 데이터, 경제 데이터, 인구조사 데이터 등 공적으로 사용 가능한 데이터 소스의 수가 증가하고 있다. 이처럼 데이터가 급증하고 있음에도, 정부나 상업적 조직이 개인의 프라이버시를 유지하는 것은 불가능하지는 않을지라도 매우 어려운 일이다.

1.3 퍼즐 조각에서부터 완성된 그림까지: 미래는 바로 지금이다

데이터의 양이 증가한 만큼, 데이터를 살펴보고 분석하는 능력도 함께 발달했다. 이전에는 우리에게 관한 작은 정보 조각들이 온라인과 오프라인에 저장되는 수준이었다면, 지금은 우리가 누구인지까지 완전히 맞출 수 있을 정도다.

역사적으로 보면 데이터 마이닝(data mining)과 예측 모델링, 그리고 탐구 분석학 분야에서 두 가지 문제가 있었다. 충분한 데이터를 저장할 수 없다는 것이 그 첫 번째요, 데이터 처리를 위한 컴퓨팅 파워에 돈이 많이 든다는 점이 두 번째 문제다. 오늘날에는 저장 장치와 데이터 처리를 위한 파워에 드는 비용이 기하급수적으로 줄어들고 있으며, 이러한 상황은 지속될 것으로 보인다. 그와 동시에 우리 각각에 대해 디지털 포맷으로 사용 가능한 데이터가 전례 없이 집적되고 있다. 덕분에 조직(이들테면 정부기관)이 한 개인에 대한 정보를 쉽게 찾을 수 있을 뿐만 아니라, 분석 모델을 사용하여 그의 미래 행동까지도 쉽게 예측할 수 있다.

개인의 정보가 어느 때보다 더 많이 공개되어 있으며 이 정보들은 개개인의 구매, 정치 성향, 또는 범죄 행위 등 모든 종류의 행동을 예측하는 데 사용될 수 있다. 이와 같은 정보들은 또한 일상적으로, 억압적인 정권이나 시민에게 치명적일 수 있는 국내외 안보에 잠재적인 위협을 식별하는 프로파일을 만드는 데, 죄가 있든 없든 높은 위험도 프로파일의 예측 모델과 매칭시킴으로써 사용되고 있다.

1.4 아주 위험한 적과 같은 광고

행동기반 광고^{Behavioral advertising}는 정말 프라이버시 관점에서 ‘아주 위험한 적’일까? 이런 개념이 전혀 새로운 게 아니라는 점은 분명하다. 이는 단순히 당신의 행동을 기반으로, 당신이 어떤 서비스나 제품의 구매에 관심을 보일지를 예측하는 방법이다.

디지털 시대 이전에는 구매 행위 분석에 특화된 AC닐슨^{AC Nielsen} 같은 회사가 있었고, 소득 수준과 성별, 결혼 여부, 구매 습성 등을 기준으로 분류한 고객리스트를 ‘임대’해주는 회사도 있었다. 영업 방법으로는 우체통에 온갖 종류의 홍보우편물을 채워 넣거나, 한 시간이 멀다 하고 무언가를 사라거나 팔라고 전화를 걸어대는 것이 고작이었다. 그런 방법을 쓸 수 있었던 이유는 아마도 우리가 은행, 신용카드 회사, 슈퍼마켓 등에 자신의 정보를 제공했거나, 잡지를 정기구독하고 있었기 때문이었을 것이다. 그러나 그렇게 얻은 정보는 다소 맹목적이었다. 빌린 사람(광고를 구매한 사업자나 조직)은 특정 그룹의 일원으로 당신의 이름이나 주소의 정보를 얻었지만, 당신이 반응할 때까지(어쩌면 반응하지 않을 때까지) 그 정보를 갖지는 못했다. 당신이 어떤 반응을 하고 나서야 비로소 당신은 그 회사의 메일링 리스트에 들어가고, 당신에 대한 고유의 이력을 쌓아나갈 수 있었다. 또한 그렇게 된다고 해도, 특정 회사나 조직의 행동에 따라 여러 가지 고객 데이터베이스에 속하게 되거나 여러 가지 버전의 이력이 존재하게 되었다.

인터넷 시대가 열린 이후에는, 가령 내가 웹 사이트를 통해 ‘하와이를 좋아한다(목표 행동)’는 사실을 보여주고 있다면, 나는 아마 웹 서핑을 하면서 하와이 여행에 관한 광고를 보게 될 것이다. 알래스카를 사랑하는 사람은 알래스카 여행에 관한 광고를 보게 될 것이다. 이것은 온라인 광고를 그냥 좀 더 개인화시킨 버전이다. 우리의 행동이 추적당하고 우리의 행동을 기반으로 여러 가지 추정을 하고 있기 때문에, 어디를 가고 무엇을 하는지에 따라 개인화된 광고에 노출되는 것이다. 광고회

사는 이러한 모델을 좋아한다. 보다 흥미를 가질 독자들을 대상으로, 그들에게 더 밀접한 광고로써 다가갈 수 있기 때문이다. 결과적으로 이는 물건을 더 많이 팔 수 있다는 것을 의미한다.

디지털 시대 이전과 지금이 다른 점은, 당신이 하는 모든 것들이 디지털 방식으로 캡처될 수 있으며 분석되어 당신을 훑아낼 수 있다는 점이다. 구글은 온라인에서의 사람들의 행동, 인구통계를 추적하며 광고 쿠키¹⁷에 관심을 갖고 있다. 많은 회사가 당신의 행동을 추적하고 있고(주로 우리가 알게 모르게 PC나 스마트폰 등의 전자 기기에 설치되도록 허락한 쿠키를 통해서), 엑셀레이트¹⁸ 같은 수많은 회사들이 당신의 정보를 팔고 있다. 그러나 대개의 경우 이런 정보들이 특별히 당신을 식별하지는 않는다. 그보다는 당신을 당신과 유사한 인구통계 자료와 관심을 가진 사람들로 분류해서 하나의 그룹에 넣고, 광고를 위해(물론 온라인 광고를 말한다) 그 그룹의 정보를 누군가에게 ‘임대’한다.

이력 정보가 여러 개이던 과거와는 달리, 당신이 누구이고 무엇을 하는지 훨씬 잘 이해할 수 있도록 이러한 그룹들을 조합하는 것은 매우 쉬운 일이다. 예를 들어, 스포키오¹⁹는 공개적으로 사용 가능한 정보를 전화번호부, 소셜 네트워크, 마케팅 설문, 부동산 목록, 상업용 웹 사이트 그리고 정부기관에서 긁어 모은다. 이 사이트에서 당신의 이름을 검색해보라. 당신이 어디에 사는지(구글 지도상의 위치), 집을 얼마에 구매했는지, 재산세는 얼마나 납부하는지(정부 데이터 소스를 통해), 집에 몇 명이 살고 있는지(인구조사 데이터를 통해)를 비롯해 당신의 모든 휴대전화 번호(온라인 전화번호부를 통해), 이전 주소와 그 집들의 가격까지 수집되어 있는 걸 알 수 있을 것이다. (소셜 미디어에 정보를 얼마나 공개했는지에 따라 다르겠

17 **역자주** 특정 웹 사이트에 방문했을 때 만들어지는 정보를 담은 파일, 방문 기록, 비밀번호 등을 사용자 컴퓨터에 기록하였다가 다음에 방문할 때 되살린다.

18 <http://exelate.com/new/>

19 **역자주** 신상정보를 모집하는 사이트. <http://www.spokeo.com/>

지만) 가까운 친구나 가족 외에는 공개하고 싶지 않은 당신에 관한 많은 정보가 이처럼 정확하게 수집되고 있다는 것이 놀랍지 않은가. 이런 정보의 대부분은 사실 인터넷에 접속하지 않고도 수집할 수는 있다. 하지만 그러기 위해서는 시간을 들여 다양한 기관을 방문해서 양식에 맞게 서류를 작성하고, 수수료를 내는 엄청난 노력을 들여야 가능하다. 과거에 그러던 것이 오늘날에는 누구든 이름을 입력하고 ‘확인’ 버튼을 클릭하기만 하면, 이 모든 것을 얻을 수 있다.

그리고 이는 더 이상 쿠키에 국한한 문제가 아니다. 예를 들어, 당신의 개인 정보를 포함한 공공 데이터가 스크랩될 수 있고(‘웹 스크래핑web scraping’이라고도 한다), 수집되어 분석될 수 있다. 또한, ‘위치 마케팅’이라는 비교적 새로운 개념도 있다. 위치 마케팅은 당신의 위치 정보(휴대전화의 GPS 칩을 통해 가능하다)에 기반하여 광고를 제공하는데, 가령 GPS상 당신의 위치가 특정한 가게 근처일 때, 그 가게에서 특별한 광고나 쿠폰을 받는 식이다.

관점에 따라 공적 또는 사적인 소스에서 수집할 수 있는 많은 양의 데이터가 거슬릴 수 있고, 단순히 디지털 시대를 살아가기 위해 치르는 대가로 여길 수도 있다. 어쨌든 당신이 이용하는 사이트(페이스북, 트위터, 링크드인, 구글, 포스퀘어 등)는 그들이 영속적으로 존재할 수 있도록 보장하는 비즈니스 모델이 필요하다. 방문하는 사이트와 당신이 맺는 암묵적인 거래는 다음과 같다. “내가 얻는 가치만큼 나도 어떤 가치를 돌려주겠다. 그 가치가 우리의 개인 정보고, 그 정보가 정기적으로 광고주에게 임대되고 있다. 보다 수익성 높은 광고를 보다 쉽게 할 수 있는 많은 정보가 있기 때문에 이제 개인 정보는 황금보다 소중하다.”

여기서 한 가지 짚고 넘어가야 할 중요한 부분이 있다. 개념상 소비자인 우리가 이러한 사실을 인지하고 있다면, 행동기반 광고는 도덕적으로 잘못된 것이 아니다. 게다가 개인 정보가 오직 광고 목적으로만 수집되고 사용된다면 그 영향은 무해한 편이다. 프라이버시 논란은 행동기반 광고에 대한 것이 아니라, 우리의 정보가 캐

내지고 사용될 수 있는, 유해한 다른 모든 방법들에 대한 것이다. 우리가 소비자로서 데이터 프라이버시를 계속 미국 증권협회 공정거래규칙과 관련 지어 생각한다면, 매우 중요한 사안을 잊고 있는 것이다. 우리의 데이터를 누가 사용하는지, 왜 사용하는지, 그리고 누가 우리를 주시하는지 전혀 모르는 상태에서 어떻게 프라이버시 침해로부터 우리 스스로를 보호할 것인지를 말이다.

1.5 세계를 장악한 빅브라더와 빅데이터

정부는 각종 범죄와 테러에 대항하기 위해 디지털 흔적을 캡처하고 분석하는 데 점점 더 많이 투자하고 있다. 그런 한편, 사람들이 프라이버시 이슈를 다룰 때 프라이버시 그 자체보다 보안을 우선시하도록 자극하는 도화선 같은 말들에 집중하고 있다. 어쨌든 우리의 생활 자체에 위기가 닥쳤는데, 어떻게 개인의 프라이버시 보호만을 주장할 수 있겠는가.

영국 정부는 자국민과 외국인 모두를 추적하기 위해 디지털 비디오 기술을 사용한다. 그들은 영국 내 185만 대 이상의 CCTV카메라를 설치했는데,²⁰ 인구에 비례해 보면 32명당 한 대 꼴이다. 런던을 가로질러 걷는 사람은 하루에 수백 번 이상 카메라에 잡힐 것이다. 영국 당국은 후드가 달린 운동복을 입는 것을 금지하여 감시를 더 쉽게 하는 방안²¹과, 나아가 범죄 직전의 행동을 식별해서 범죄가 일어나기 전에 경찰관이 출동할 수 있도록 인공지능 프로그램을 사용하는 방안도 고려하고 있다.²²

미국의 많은 법 집행기관들은 데이터 수집과 분석 기술에 많은 부분 의존하고 있다. 뉴욕 시 경찰은 소위 ‘정지 신체 수색권²³’을 개인에게 적용하여 그들의 이름,

20 <http://www.securitynewsdesk.com/2011/03/01/only-1-8-million-cameras-in-uk-claims-acpo-lead-on-cctv/>

21 <http://www.express.co.uk/news/uk/39622/Ban-the-hood-for-good>

22 <http://www.pcpro.co.uk/news/208452/ai-could-power-next-gen-cctv-cameras>

23 **역자주** 우리나라의 불심검문과 비슷한 개념 <http://www.nyclu.org/news/nypd-stopped-record-number-of-innocent-new-yorkers-2010-new-stop-and-frisk-numbers-show>

신체 특징, 신분증, 그리고 동반자의 이름을 중앙 데이터베이스에 입력한다. 경찰관이 ‘정지’시킬 정도가 아닌 범죄적 행위는 굳이 관찰할 필요가 없음에도 2010년에만 59만 명이 넘는 사람들을 대상으로 이런 작업이 수행되었고, 그들 대부분은 흑인과 히스패닉이었다. 이처럼 노골적으로 차별하는 ‘정지 신체 수색권’을 통해 개인을 구속함으로써 데이터베이스를 유지하는 법 집행은, 2009년에 주 법률이 불법으로 판결하여 더는 허용되지 않는다. 그러나 ‘정지 신체 수색권’(그리고 어떤 범죄 혐의도 없는 개인들을 대상으로 한 CCTV 영상 데이터베이스와 다른 데이터베이스도)은 뉴욕 시의 데이터 집중 분석을 위한 기관인 ‘실시간 범죄 센터’²⁴Real Time Crime Center²⁴에서 여전히 중요한 역할을 하고 있다.

감시 기술은 미국 전역으로 퍼져 나가고 있다. CCTV카메라는 고속도로망에 설치되어 교통량을 관찰하고 있으며 신호등은 정지 신호 위반을 감시한다. 우편으로 교통 위반 딱지를 받는 것은 이제 예삿일이다. 실행에 논란의 여지가 있을뿐더러 종종 헌법적 근거에 이의가 제기됨에도 계속 유지되고 있다. 비행기의 블랙박스과 유사한 ‘디지털 사건 기록기(차에 설치되는 블랙박스라고도 한다)’²⁵는 사고 발생 시 법 집행 기관이 누구의 과실인지 판별하는 데 사용된다. 렌트카 업체는 과속하거나 비포장 도로를 달렸을 때 벌금을 물리기 위해, 이와 유사한 기술에 GPS 기능을 추가하여 사용한다.

이런 정황으로 볼 때 미국 정부는 디지털 시대의 프라이버시의 유지에 대해 다른 견해를 가진 것 같다. 국제적으로는 프라이버시를 민주화의 힘처럼 여긴다. 예를 들어, 정부는 트위터나 페이스북 같은 소셜 네트워크 툴이 쉽게 파괴되지 않고 안전하도록 보장하기 위해 기술 제공자에게 보조금을 지급한다. 이런 툴들은 시리아나 튀니지, 이란 같은 국가의 친민주주의 시위대에 의해 더욱 효과적으로 사용될

24 <http://www.nytimes.com/2010/02/18/nyregion/18tattoo.html>

25 <http://www.slate.com/id/2087207/>

수 있다. 물론 이들 국가의 정부도 체제를 위협하는 세력의 제거를 목표로 같은 툴들을 사용할 수 있고, 불안한 상황이 유지될 경우 모든 접속을 차단할 수 있다고 알려져 있다.

국가 안보와 관련된 문제에 대해서라면, 미국 정부는 개인 정보에 지금보다 더 많이 접속하기를 원한다. 예를 들어 미국 사법부는 최근 미국의 안전과 안보를 지속하기 위해 ‘전자통신프라이버시법안²⁶Electronic Communications Privacy Act’의 조문을 유지해야 한다고 주장했다. 이 조문에 따르면, 지메일^{Gmail}이나 핫메일^{hotmail} 같은 호스트 업체에 저장된 개인 메일은 6개월 이상 되었을 때에 한해, 권한 없이도 열람할 수 있도록 규정한다. 법률의 이름과는 완전히 어긋난 조문이지 않은가? 소위 애국자법^{Patriot Act}을 통해서 법 집행기관은 전반적인 감시를 위한 권한²⁶을 특별법원에 요구할 수 있는데, 그 기준은 합당한 이유를 요구하는 일반적 기준에 비해 상당히 낮은 수준이다. 이 법률을 적용하면 한 개인에 대한 모든 기록을 열람할 수 있는데, 요청 자체가 비밀로 취급되기 때문에 해당 개인은 그 사실을 알 수조차 없다.

미국 헌법상 프라이버시에 관한 명확한 언급은 없지만, 대법원에서 프라이버시의 반영적 권리²⁷로서 권리장전에 몇 가지 사항을 개정해왔다. 이는 법적으로 논란이 있는 부분이고, 우리는 변호사가 아니므로 ‘실세계(디지털이 아닌 현실세계)’에서 무엇이 사적이고 무엇이 아닌지에 대한 법적 정의에 대해서는 중립적인 입장을 취해야 하겠지만, 그래도 디지털 세계의 그것과 비교해보면 아주 명확하다. 다시 말해, 디지털 세계에서나 실세계에서 우리의 ‘프라이버시 권리’는 끊임없이 변해간다. 다만 ‘현실세계’에는 ‘자리를 잡은 법(선례 구속성의 원리²⁸)’의 합법적 기준에 접근한 전례가 존재한다. 그러나 그것을 운영하는 기술 등에 관련한 디지털 세계에는, 프라이버시 법에 관한 자리를 잡은 법안이 없다.

26 <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1198/1118>

27 <http://plato.stanford.edu/entries/privacy/>

28 http://en.wikipedia.org/wiki/stare_decisis

이것은 우리가 당면한 근본적인 질문이다. 디지털 시대에 우리는 정부로부터 감시 받지 않을 권리를 가지고 있을까? 있다면, 어디에 있는가? 인터넷상에서, 서재에서, 공공장소에서, 사사로운 업무 중에, 고속도로 위에서, 또는 정부를 반대하는 평화적인 시위에서? 1759년에 벤자민 프랭클린^{Benjamin Franklin}은 이렇게 말했다. “일 순간의 안전을 위해 필수적인 자유를 포기하는 자는 자유도 안전도 누릴 자격이 없다.” 프라이버시 대 보안의 문제는 언제나 매우 어려운 주제다. 더욱이 개인의 사적인 디지털 데이터에 쉽게 접속하고 그 정보를 추적할 수 있다는 사실은, 테러가 빈번하고 대량살상무기가 존재하는 시대에 이 문제를 더욱 복잡하게 만든다.

1.6 프라이버시 대 보안과 안전, 그 갈림길에서

페이스북의 마크 주커버그가 말했듯이, 디지털 시대에서 프라이버시는 시대에 뒤쳐진 주제일까? 결론적으로 말하자면, 당신이 잘못된 것이 없다면 걱정할 게 없다. 물론 당신이 인종적인 관심이나 종교적인 관심을 지닌 누군가에게 그렇게 말한다면 그의 반응에 놀랄 수도 있다. 우리는 갈림길에 서 있다. 우리는 프라이버시를 얼마나 기꺼이 포기할 수 있는가? 우리는 얼마나 투명하게 지내기를 원하는가? 우리는 정부가 우리를 감시하는 것을 어느 정도까지 수용할 수 있는가? 범죄나 테러리즘에 관하여 우리의 개인 정보 보호를 위한 대가로서 어느 정도의 위험을 감수할 수 있는가? 그러한 위험을 어떻게 측정할 것인가? 보다 안전한 프라이버시를 위해 포기할 수 있는 정확한 수치를 우리는 어떻게 알까?

휴대전화에 저장된 사진을 온라인을 통해 공유할 때, 사진이 촬영된 위치를 정확히 알려주는 GPS 정보가 포함될 가능성이 있다. 그것이 만약 학교에서 찍은 자녀 사진이라면 당신은 학교의 위치가 알려지는 것을 원치 않을 것이다. 또한 당신은 가령 커뮤니티 사이트에 접속해서 가족 구성원이 얼마나 아픈지 공유할 수도 있다. 이제 의료보험에 대해 알아보면 보험회사는 당신도 모르게 그 정보를 가지고 있을 것이다. 휴대전화에 내장된 GPS 추적 기능을 꺼버리면 당신의 위치를 알 수 없

을 것 같은가? 법 집행기관은 여전히 당신의 위치를 찾을 수 있다. 구글, 이베이 그리고 다른 수많은 기업을 이용해 일 년간 국민의 정보를 저장하고자 하는 프랑스에 산다면, 당신을 조사해야 하는 프랑스 기관이 그것에 접속하려 할 것이다. 트위터를 사용한다면 멘션을 작성한 위치 또한 추적당할 수 있다. 페이스북을 이용하여 혁명을 피하고자 한다면, 당신의 시위 대상이 되는 정부 역시 페이스북을 통하여 당신을 감시할 수 있다.

이러한 정보의 수집과 추적은 당신의 승인을 받고 이루어진다. 그러나 많은 기업과 기관에서는 여러 가지 방법으로 그 승인을 침해한다. 옵트인²⁹ 방식인 것으로 간주해서 어쩔 수 없이 직접 옵트아웃³⁰ 방식³⁰해야 한다는지, 알지 못하는 사이 쿠키를 데스크톱에 저장해 놓거나, 정보를 모으는 수상한 실험을 하고, 정보를 공유하지 말라고 요구함에도 공유하는 식이다. 기술적 발전으로 말미암아 훔쳐보기는 쉬어진 데 반해, 우리 자신을 보호하는 데 필요한 것들을 갖추기는 어려워졌다.

우리를 보호하는 것이 정부의 역할이라고 생각하는가? 잠시 이 점을 생각해보기 바란다. 미국에는 개인의 프라이버시를 보호하는 내용의 30개 이상의 연방 법령과 100개 이상의 주 법령이 있다. 그러한 규정은 단편적일뿐더러, 기업이 자기검열을 통해 당신을 침범하지 않도록 보호하기 위해 작성되었다. 계류 중인 인터넷 권리 장전도 있고, 텔레마케터를 통제하는 Do-Not-Call 리스트³¹와 유사한 Do-Not-Track시스템³²도 있다. 또한 TRUSTe³³처럼 당신의 프라이버시를 보호하기 위해

29 **역자주**_옵트아웃과 반대로, 사전 동의 없이 일단 무조건 동의한 것으로 처리한 후, 나중에 사용자가 거부 의사를 밝혀야만 중단시키는 방식

30 **역자주**_신청을 통해 사전 동의를 한 사람에게만 어떤 제도나 서비스가 적용되는 방식

31 **역자주**_텔레마케터의 전화를 원치 않는 전화번호나 웹 사이트를 미국 연방정부가 운영하는 데이터베이스에 등록하여 관리하는 것

32 **역자주**_크롬 등의 인터넷 브라우저가 지원하는 기능으로, 웹 사이트에서 개인 정보를 활용하는 것을 개인이 선택할 수 있게 한다. '사용자 추적 방지'라고도 부른다.

33 <http://www.truste.com/>

만들어진 소비자 프라이버시 기관과 활동 그룹, 그리고 기업도 있다. 인터넷은 세계화되었지만 프라이버시 이슈는 그렇지 못해서 프라이버시 보호법이나 규제 조치, 단체들은 국가마다 다르다.

우리는 복잡한 세계에서 살고 있다. 프라이버시 당사자와, 단속 기관, 이해 관계자들이 공존한다. 이들은 오늘날의 프라이버시 상태와 일어나는 일들에 대해 당당하게 맞설지 아니면 두려워해야 할지 장황하게 늘어놓는다. 우리가 잃어가고 있는 것은 바로 ‘프라이버시 권리’의 의미다.

- 온라인에서 공유하는 개인 정보에 대하여 우리는 어떤 가정을 해야 하는가?
- 누가 우리의 정보를 가지고 있으며 그들은 그 정보로 무엇을 하는가?
- 미국과 그 외 국가에서 개인을 보호하기 위해 어떤 규정을 정해두는가?
- 정보 프라이버시에 대해 요구하고 관련 법을 만들기 위해 노력하는 이들은 누구인가?
- 민주주의 사회에서 디지털 정보를 사용해도 되는 적법한 정부란 무엇인가?
- 소비자로서 이 모든 것을 받아들이기 위해서는 어떠한 자세를 갖추어야 하는가?

1597년에 프란시스 베이컨 경은 “아는 것이 힘이다”라고 말했다. 이 말은 그때나 지금이나 참이다. 빅데이터 시대에 살면서, 프라이버시에 대해 알면 알수록 데이터 프라이버시에 관한 정책, 표준, 규정을 만드는 데 영향을 끼칠 수 있다. 이 책은 광고에 관한 토론이 아니다! 날이 갈수록 투명해지고 위험해지는 세계에서 우리가 프라이버시, 보안, 안전의 균형을 어떻게 잡을지에 관한 심도 깊은 토론이다.

1.7 참고 문헌

1. Kim Zetter, “Feds ‘Pinged’ Sprint GPS Data 8 Million Times Over a Year” Wired, December 1, 2009
2. Cameron Chapman, “The History of the Internet in a Nutshell” Six Revisions, November 15, 2009
3. Wikipedia, “Internet”
4. Wikipedia, “AOL Instant Messenger”
5. Wikipedia, “SMS (Short Message Service)”
6. Berkman Center for Internet and Society, Harvard University, “A History of Digital Data Creation”
7. The Radicati Group, Inc., “Key Statistics for Email, Instant Messaging, Social Networking and Wireless Email” April 19, 2010
8. Pew Research Center, “Pew Internet and American Life Project, Internet Trend Data”
9. Joseph Turow, Jennifer King, Chris Jay Hoofnagle, Amy Bleakley, Michael Hennessy, “Americans Reject Tailored Advertising and Three Activities that Enable It” September 29, 2009
10. Opera Software Press Release, “Who’s Watching You” January 28, 2011
11. Michael Rappa, “Business Models on the Web” Managing the Digital Enterprise, January 17, 2010
12. Google, “Google Announces Fourth Quarter and Fiscal Year 2010 Results and Management Changes” January 20, 2011
13. IDC Research, “The Digital Universe Decade” May 2010

14. NOAA (National Oceanic and Atmospheric Administration), "[Tsunamis May Telegraph Their Presence](#)" January 19, 2010
15. Adam Singer, "[49 Amazing Social Media, Web 2.0, and Internet Stats](#)" The Future Buzz, January 12, 2009
16. Facebook, "[Press Room Statistics](#)"
17. Nathan Wolfe, Lucky Gunasekara, and Zachary Bogue, "[Crunching Digital Data Can Help the World](#)" CNN, February 2, 2011
18. Terri Wells, "[Website Marketing: How and Why Behavioral Advertising Works](#)" November 1, 2006
19. Matt Drake, "[Ban the Hood for Good](#)" EXPRESS.co.uk, March 30, 2009
20. Stuart Turton, "[AI Could Power Next-gen CCTV Cameras](#)" PC PRO, June 25, 2008
21. New York Civil Liberties Union, "[NYPD Stopped Record Number of Innocent New Yorkers in 2010, New Stop-and-Frisk Numbers Show](#)" February 23, 2011
22. Michael S. Schmidt, "[Have a Tattoo or Walk With a Limp? The Police May Know](#)" New York Times, February 17, 2010
23. Wikipedia, "[Electronic Communications Privacy Act](#)"
24. Mary Minow, "The USA PATRIOT Act and Patron Privacy on Library Internet Terminals" LLRX, February 15, 2002
25. Joan Starr, "Libraries and National Security: An Historical View" First Monday, December 6, 2004
26. Stanford Encyclopedia of Philosophy, "[Privacy](#)" September 18, 2006

27. Wikipedia, “[Confrontation Clause](#)”
28. Wikipedia, “[United States Bill of Rights](#)”
29. Wikipedia, “[NSA Warrantless Surveillance Controversy](#)”
30. Wikipedia, “[President’s Surveillance Program](#)”
31. [Unclassified Report on the President’s Surveillance Program](#), July 10, 2009
32. USA Today, “[NSA Has Massive Database of Americans’ Phone Calls](#)” May 11, 2006
33. Philadelphia Independent Media Center, “[Why We Should Be VERY WORRIED about How Bradley Manning Is Being Treated](#)” March 15, 2011
34. Mobile Marketer, “[Location-based Marketing Can Increase Average Order Value, Frequency, Loyalty](#)” Dan Butcher, March 29, 2011
35. Fast Company, “[Google, eBay, and Facebook Take on France Over User Privacy](#)” Austin Carr, April 5, 2011
36. Managing the Digital Universe, “[Data Privacy](#)” Michael Rappa, January 17, 2010
37. The Wall Street Journal, “[Proposed Bill Would Put Curbs on Data Gathering](#)” Julia Angwin, March 10, 2011
38. ReadWriteWeb, “[What Twitter’s New Geolocation Makes Possible](#)” Marshall Kirkpatrick, November 19, 2009