



# The *Personalized* and the *Personal*

Socially responsible innovation through big data

BY JAMES GUSZCZA, DAVID SCHWEIDEL, AND SHANTANU DUTTA  
> ILLUSTRATION BY LIVIA CIVES

*“The best minds of my generation are thinking about how to make people click ads. That sucks.”*

—Jeff Hammerbacher, Cloudera founder

인터넷은 우리 삶의 필수불가결한 부분이 되었으나, 오늘날 도처에서 일어나는 데이터 수집과 데이터 분석기술의 확대로 사회 전체가 불편함을 느끼고 있다. 이 같은 불편함에 대한 언급은 기자, 정치인, 소설가에 국한되지 않고, 업계 당사자인 네트워크 이론가 또한 ‘빅 데이터의 윤리적 사용’을 주장하고 있다. 이는 유명 데이터 과학자 사이에 점점 더 퍼져나가는 “데이터 회의론 Data skepticism”적인 태도와 관련이 있다. 데이터 회의론이란 데이터 과학의 본질적 가치에 대한 회의가 아니라, (의도적이든 아니든) 잘못된 의사결정이나 비윤리적인 실무 관행을 위한 데이터 또는 애널리틱스 모델의 미심쩍은 사용에 대한 회의론이다.

데이터 회의론이 점점 부상하는 이유는 데이터의 방대함에 대한 과장 때문이 아니라, 많은 이들이 공격적이거나 “소름 끼친다”고 느끼는 심리적, 행동적, 사회적 콘텐츠를 애널리틱스가 다루고 있기 때문이다. 이 같은 분위기는 오히려 데이터 중심 기업들로 하여금 자사의 데이터 전략이 가지는 윤리적, 사회적 함의를 좀 더 면밀하게 평가할 동기를 제공하여, 지속 가능한 성장을 위한 혁신 전략 구축으로 이어질 수 있다. 그러나 이를 위해서는 데이터의 어떤 사용을 많은 사람들이 왜 소름 끼친다고 여기는지 정면으로 맞설 필요가 있다.

이 사안은 빅 데이터보다는 행동 데이터에 해당된다. 전례 없이 많은 일상 정보가 디지털로 파악되고, 개인적 특성을 추론하는데 사용할 수 있게 되었다. 예측할 수 없는 다양한 데이터를 조합해 구성한 개인의 초상은 우리의 외적 인격뿐만 아니라 개인적인 자기 인식과도 불일치할 수 있다. 많은 이들이 이러한 점을 불편하게 여기는 것은 이해할 만하다. 그러나 나쁜 면만 있는 것은 아니다.

현대 행동과학 분야의 연구는 중요 의사결정과 목표 달성 과정에서 사람들이 겪는 어려움에 상당 부분 집중되어 있다. 이러한 연구는 보다 나은 의사결정을 가능하게 하는 혁신적인 데이터 상품과 서비스의 아이디어 원천이 될 수 있다. 비즈니스에서의 데이터 분석을 최근에 형성된 “우리 대 그들”의 대결 구도에 맞출 필요가 없으며, 반대로 행동과학에서 얻은 통찰뿐 아니라 윤리적 사고로 활력과 동기를 가지게 되면, 데이터 애널리틱스를 통해 선한 의도로 성공할 수 있는 21세기적인 접근법을 얻을 수 있을 것이다.



### LET'S GIVE CREDIT SOME THEORY

1960년대 비즈니스 애널리틱스 혁명의 시초가 된 개인 신용 데이터는 행동에 대한 “디지털 흔적”의 초기 사례이다. 수십 가지의 약한 예측성 데이터 요소들을 결합해 더 강력한 예측력을 가지게 만드는 예측 모델을 사용하는 신용등급 평가 알고리즘은 대출 심사역보다 훨씬 정확하고 경제적인 것으로 밝혀 졌다. 오늘날 대부분의 개인대출은 신용등급을 사용해 승인되고 금리가 결정된다.

30년이 지난 후, 개인보험업계는 자동차보험 가입자의 사고 발생 가능성과 주택 소유자 종합보험 가입자의 보험 청구 가능성을 개인 신용 정보를 통해 예측 가능하다는 사실을 발견했다. 많은 사람들이 이 연관성을 놀라다고 생각할 것이며, 불편해할 수도 있다. 비용 납부 행동이 왜 자동차 사고 가능성과 관련이 있는가? 전자와 후자를 연결하는 인과관계가 있을까?

이 주제를 연구한 University of Texas-Austin의 Patrick Brockett, Linda Golden 교수의 결론에 따르면, 신용평가 점수는 개인의 재무 행동을 요약해 주는데 관찰 불가능하지만 예측력이 강한 무언가에 대한 프록시 역할을 한다. 대출 및 상환 같은 종류의 행동은 다른 종류의 행동(방어적 또는 공격적 운전 등)에 영향을 미치는 잠재적 행동특성을 보여주는 징후일 가능성이 높다. 보다 은유적으로 말하면, 신용평가 점수는 분명한 비즈니스 연관성이 있는 예상치 못한 마음의 창을 제공해준다.

### KNOWING ME, KNOWING YOU

신용평가 점수뿐 아니라, 정치권에도 사례가 있다. 2012년 오바마 대통령의 재선 캠페인에서는 인구 정보뿐 아니라 과거 투표, 생활방식 등을 포함하는 방대한 데이터베이스를 구축하여 행동과학자 팀이

분석한 결과 유권자의 “실득가능성” 측면에서 각 유권자를 분류할 수 있었다.

더욱 놀라운 사례가 신생 분야인 신경경제학(Neuroeconomics)에서 나왔다. 금융 트레이더들의 약지와 검지 길이의 비율(“2D:4D”)이 그들의 업계 종사기간 뿐 아니라 장기 수익성을 예측하는 프록시 역할을 한다고 밝혀진 것이다. 2D:4D의 값은 태아일 때 테스토스테론 남성호르몬에 대한 노출 정도와 관련이 있는데, 이는 심장 질병, 당뇨병, 알콜중독, 학교 성적, 비즈니스 리더십/혁신 정도와 연관관계가 있다. 2D:4D의 측정이 빅 데이터에 해당하지는 않지만, 일련의 행동에 대한 예측력은 높다.

인터넷 시대를 사는 덕에 우리는 매우 사적인 개인 특징을 엿볼 수 있는 디지털 프록시 구축 능력을 가지게 되었다. 페이스북 상의 “좋아요 Likes”버튼을 누른 데이터를 추출하여 진행한 최근 연구 결과에서는 인종(95%), 종교(82%), 성적 취향(88%), 정치 성향(85%) 등을 높은 정확도로 예측하였다.

각각의 개인 디지털 흔적의 정보량은 매우 적지만 복합적인 측정방법을 적용하면 사생활, 즉 개인 특징과 심리적 기질 등을 집합적으로 보여주는 강력한 지표가 된다. 행동 예측은 마케팅, 보험, 보건, 정부, 교육 등의 많은 영역에서 정말 중요하다. 위의 사례가 함의하는 것은 예측적인 행동 패턴이 예상치 않았던 곳에서 발견될 수 있다는 사실이다. 그리고 우리 일상활동이 더욱 전자화되어 기록될수록 데이터 기반의 행동 예측 기회는 크게 증가할 것이다.

## O BRAVE NEW WORLD THAT HAS SUCH ALGORITHMS IN IT

과거의 모든 구매, 대출, 결제 활동은 대부분 기록 없이 이루어졌으며, 총 회계 통계에만 반영되었다. IT 사용이 늘어나면서 모든 거래 기록이 전자적으로 기록되기 시작했다. 개인의 재무 거래가 디지털 흔적을 남기기 시작한 것이다. 처음에는 회계 부기의 편의 목적으로 시작되었지만, 사업가들은 마침내 과거의 행동이 미래 행동을 예측한다는 사실을 깨달았다. 수십 년 동안 이러한 예측은 데이터가 생성되는 동일 영역인 개인 금융에서 중점적으로 이루어졌는데, 이제는 이 데이터가 다른 영역에서도 놀랄 만한 효과가 있음이 밝혀졌다.

처음에는 온라인 상에서 수집되는 정보가 개인 금융 행위 뿐이었지만, 요즘에는 TV 시청, 동영상 감상, 소셜 미디어 활동, 쇼핑, 운전 행동양식 정보, 심지어 “자기 측정(Quantified self)”이라는 흐름에 따라 수면 행동, 다이어트 등 우리의 일상 생활까지 디지털화 되어 측정, 수집되고 있다.

성실한 임차인 선별, 이혼 부모의 양육비 지급 체불 가능성 등 신용평가 점수의 적용 가능 범위가 제한이 없는 것처럼, 우리가 남기는 수많은 디지털 흔적이 상상하기 힘든 방식, 아마도 받아들이기 어려운 방식으로 아주 유용하다는 사실이 입증될 것이다. 일례로, 정부가 모든 형태의 디지털 흔적을 사용해서 시민들의 범죄행위 가능성을 사전 예측하는 것이 바람직하다고 여기게 될까? 이 시나리오는 Philip K. Dick의 동명 소설에 기초한 영화 마이너리티 리포트 Minority Report로 만들어 졌다. Dick의 허구로 만들어낸 “예언자 Pre-cogs”의 영적인 힘은 불필요하다는 것을 이제 우리는 안다. 개인행동 예측은 다양한 디지털 흔적을 축적하고 통계적으로 분석하는 것으로 충분하기 때문이다.

데이터 분석이 과학적으로 타당하지만, 사회적으로 바람직한지는 전적으로 다른 문제이다. 즉, 데이터를 다루고 이용하는 데 있어 기업 윤리가 중요한 것이다.

### 'TIS NEW TO THEE

민간 분야 또는 공공 분야의 행동 데이터, 소셜 네트워크 데이터, 개인 데이터 사용에 대한 거부감 정도는 사람마다 다르다. 연령별, 정치적, 철학적, 사회적 차이가 개인정보 보호에 대한 태도에 영향을 미친다. 사람들의 태도는 기술의 진화와 데이터 사용의 혁신에 따라 같이 진화할 것으로 예상된다.

이에 대한 쉬운 해결책은 없다. 한 가지 작은 제안을 하자면, 조직이 데이터의 전략적 사용을 구상할 때, 목표 대상(고객, 시민, 환자, 학생 등)이 개인 정보의 노출과 전반적인 사회적 변화라는 비용에 준하는 혜택을 돌려 받을 수 있음을 인지하고 있는지를 고려해야 한다. 마트의 회원카드, 무료 이메일 서비스 등은 데이터 제공을 대가로 혜택을 누리는 사례이며, 지금까지 규제기관과 대중의 반발이 거의 없었다. 고객 중심 관점은 또한 혁신적인 제품과 사업 모델이라는 긍정적 혜택의 원천이 될 수 있다. 아래 몇 가지 사례가 이 개념을 잘 보여준다.

**보험:** 자동차 보험사는 GPS 장치로 수집한 운전행동 데이터를 보험료 책정과 운전자 세그멘테이션 등 핵심분석 업무의 입력 자료로 사용하는 방안을 고려 중이다. 그러나 이외에도 보험 가입자에게 “데이터 상품”의 형태로 부가 가치를 제공할 수 있는데, 운전 습관에 대한 개인 보고서, 젊은 운전자의 또래 집단과의 비교 분석 보고서 등이다. 이러한 혁신적인 데이터 상품의 제공은 경쟁업체로부터 차별화, 고객호감도 구축, 사회적 책임 서비스 제공을 동시에 이룰 수 있는 경제적 전략이 될 수 있다.

**유통업체:** Tesco는 회원카드 쇼핑 데이터를 분석해서 비만, 당뇨 등의 위험을 보이는 고객을 찾아야채, 과일 등 건강 식품 쇼핑 제안을 권장하는 등의 보완상품 판촉 방법을 고려하고 있다.

**소비자 금융:** 사람들은 대부분 개인 금융 관련 의사결정을 어려워한다. 소매은행과 증권사는 데이터와 행동과학 정보를 기반으로 하여 고객의 금융 이해력과 재무 관련 의사결정 능력을 개선하는 전략을 조사할 수 있다. 의사결정의 어려움을 예측적인 “선택모델 Choice models”과 고객 세그멘테이션 분석으로 개인이나 가족의 상황에 알맞은 맞춤형 선택 메뉴를 만들어서 대처할 수 있다.

각 분야에서 데이터가 풍부한 조직은 선택을 할 수 있다. 단순히 데이터를 사용해 시장점유율을 높이고 보다 효과적으로 분석 업무를 수행하거나, 또는 이를 넘어서 보다 혁신적인 상품과 서비스를 고객에게



제공하는 방법을 연구할 수 있다. 데이터 분석 방법을 행동경제학 원칙과 결합하여 -고객의 참여 하에- 그들의 개인적 목표 달성에 도움이 되는 정보와 도구를 제공해야 한다.

## NEW BOTTLES FOR NEW WINE

행동경제학의 대전제 중 하나는 사람들이 관련 지식과 최선의 의도를 가지고 있어도 효과적인 행동에 실패한다는 것이다. 나쁜 운전습관 지속, 다이어트 실패 등의 좋은 예이다. 인간의 마음은 중요한 업무나 의문에 직면했을 때, 자연스럽게(중증 무의식적으로) 이를 미루거나 손쉬운 것으로 대체하려는 경향이 있다.

앞에서 언급된 유형의 행동과학 기반 도구를 통해 사람들은 이득을 볼 수 있다. 전통적 목표 설정과 세그먼트 기반의 성장 전략에서 나아가, 기업은 고객이 더 나은 선택을 하고 초심을 유지할 수 있도록 도움을 주는 고객 데이터 상품과 서비스를 제공할 수 있다.

정부와 기업의 빅 데이터 사용에 대한 대중의 의구심은 데이터가 가지는 힘의 원천에 기인한다. 과거에 숨겨져 왔던 개인 행동이 디지털 흔적을 남기면서 전문가나 일반인 모두 예측하기 어려운 방식으로 사용될 수 있기 때문이다. 데이터 수집자와 사용자가 공익을 우선한다는 신뢰가 없으면 대중의 의구심은 줄지 않는다.

세부적인 행동정보와 행동경제학, 선택형 구성(Choice architecture) 등을 기반으로 데이터 상품과 서비스를 만드는 것이 가능해지면, 대중의 신뢰는 오히려 늘어날 것이다. 이러한 아이디어를 수용하는 기업은 더욱 뛰어난 소비자 지향 상품 설계를 통해 경쟁업체와 차별화를 이루고, 지속적인 수익성과 성장을 누릴 수 있을 것이다. DR

*James Guszczka*는 Singapore Deloitte Analytics Institute의 선임연구원이며 Deloitte Consulting of Actuarial, Risk and Advanced Analytics 부문의 US National Predictive Analytics Lead로 재직 중이다.

*David Schweidel*은 Emory University의 Goizueta Business School의 마케팅 부교수로 재직 중이다.

*Shantanu Dutta*는 University of Southern California의 Marshall School of Business의 마케팅 담당 교수이자, Singapore Management University의 객원 교수이다.

