문항1. 빅데이터를 활용하는 목표 4가지에 대해 서술하시오.(72점)

빅데이터 1 : 사전 계획

가트너의 휴데커는 “흔한 말이지만 ‘계획하지 않으면 실패를 계획하는 것’이라는 말이 여기에도 적용된다. 규모가 작고 달성 가능하며 새로운 것을 선택하라. 제약이 있는 레거시 사용 사례를 선택하면 안 된다”고 말했다.

PwC의 모리슨은 “데이터에 대해 먼저 생각하고, 데이터가 그 조직에 유용하도록 기계가 읽을 수 있는 방식으로 조직을 모델링해야 한다”고 말했다.

빅데이터 2 : 협력

빅데이터 프로젝트에 이해당사자, 즉 프로젝트의 결과를 사용할 사람들이 배제되는 경우가 많다. 휴데커는 모든 이해당사자가 협력하면 많은 장애물을 극복할 수 있다면서 “숙련된 기술 담당자들이 상호 협력하고, 또 비즈니스 당사자와 협력하면 실천 가능한 결과를 도출할 수 있다”고 말했다.

휴데커는 빅데이터에서 성공하는 기업은 필요한 스킬에 대대적으로 투자한다는 점을 강조했다. 이런 기업 중에는 금융 서비스나 우버(Uber), 리프트(Lyft), 넷플릭스와 같이 양질의 실질적 가치를 지닌 데이터에 기업의 성패가 좌우되는 데이터 지향 기업이 가장 많다.

탈렌드의 크리스토퍼는 “데이터를 수집하고 정제하는 과정을 팀으로 수행해야 한다. 그러면 데이터의 무결성도 높아진다”고 덧붙였다.

빅데이터 3 : 집중

빅데이터 프로젝트는 크고 웅대해야 한다는 선입견을 가진 경우가 많다. 처음 무언가를 배울 때 늘 그렇듯이 성공하기 위한 최선의 방법은 작게 시작한 다음 차츰 목표와 범위를 넓히는 것이다.

휴데커는 “아주 좁게 정의할 필요가 있다. 사기 탐지, 고객 미세 구분, 밀레니얼 시장에 적합한 신제품 찾기와 같은 문제 영역을 선택해서 집중해야 한다”고 지적했다.

크리스토퍼는 “결국 원하는 통찰력 또는 디지털화할 비즈니스 프로세스가 무엇인지를 물어야 한다. 기술이 그냥 비즈니스 문제를 해결해주지는 않는다. 먼저 정의해야 한다. 데이터 레이크는 필요하지만 비즈니스에 사용되지 않을 데이터는 수집하지 말아야 한다”고 말했다.

많은 경우 이는 지나치게 비대해지면 안 된다는 것을 의미하기도 한다. PwC의 모리슨은 “연구 대상 기업을 살펴보니, 전체 비즈니스가 실행되는 핵심적인 개념과 관계는 수백 개에 불과했다. 이 점을 이해하고 나면 수백만 가지의 구분이 이 수백 가지 중요한 것의 미세한 변형에 불과하다는 것을 알게 된다. 사실 작은 변형의 상당수는 애초에 변형도 아니다. 이름과 구조, 레이블만 다를 뿐 동일한 것”이라고 설명했다.

빅데이터 4 : 레거시 폐기

수집해서 데이터 웨어하우스에 저장해 둔 테라바이트 규모의 데이터를 활용하고 싶겠지만, 사일로가 없도록 설계된 빅데이터용 스토리지 시스템에 새롭게 수집한 데이터에만 집중하는 편이 더 나을 수 있다.

컨설턴트 그린바움은 “단순히 라이선스가 있다는 이유만으로 기존 기술 인프라에 묶여 있을 필요는 없다. 복잡한 새 문제에는 복잡한 새 솔루션이 필요한 경우가 많다. 10년 전부터 사용한 오래된 툴을 고수하는 것은 바람직한 방법이 아니다. 많은 기업이 오래된 툴을 사용하며 이로 인해 프로젝트가 실패한다”고 지적했다.

모리슨은 “스스로 발목을 잡지 말아야 한다. 더 많은 사일로를 생성하는 레거시 아키텍처를 버려야 한다”고 강조했다. 또한 모리슨은 기술 업체가 복잡한 시스템 문제를 대신 해결해줄 것이라고 기대하면 안 된다면서 “수십년 동안 많은 기업이 빅데이터 문제를 해결할 방법을 돈을 주고 살 수 있다고 생각했다. 모든 빅데이터 문제는 시스템 문제다. 복잡한 시스템 변경이 필요한 경우 그 방법을 기업 스스로 만들어야 한다”고 덧붙였다

## 문항2. 스몰데이터의 정의와 필요성에 대해 서술하시오.(28점)

스몰데이터는 미국의 브랜딩 전문가인 마틴 린드스톰(Martin Lindstrom)이 제시한 개념으로, 개인의 취향과 라이프스타일 등 사소한 행동 등에서 나오는 개인화된 데이터를 의미한다.

반면 빅데이터는 컴퓨터 및 처리기술이 발달함에 따라 디지털 환경에서 생성되는 데이터로, ▷초대용량의 데이터 양(volume) ▷다양한 형태(variety) ▷빠른 생성 속도(velocity) 등을 특징으로 한다. 빅데이터가 특정 집단의 공통적인 특징을 구분해 내는 데 유리하다면, 스몰 데이터는 이용자 개개인의 취향·소비성향·기호 등을 세부적으로 파악하는 차이가 있다. 즉, 스몰데이터는 공통적인 성향이 아니라 개인의 차별화된 특성을 파악하는 데 유리하다.

이처럼 스몰데이터는 빅데이터가 포착하지 못하는 개인의 세세한 부분을 추출해 낼 수 있기 때문에, 이를 마케팅 등에 활용하면 고객의 니즈를 거의 충족시킬 수 있다는 장점이 있어 빅데이터의 보완재로 제시되고 있다.