**빅데이터의 3가지 종류**

빅데이터는 사용자가 원하는 것이 생기면 거기에 해당하는 데이터를 찾아서 분석해 제공하는 것이다. 따라서 규모(Volume), 변화 속도(Velocity), 다양성(Variety)이라는 3가지 특징, 즉 3V를 가지고 있는 데이터에서 효과적인 결과물을 보여주는데, 3가지 특징은 꼭 가져야 한다기보다 효과적이라는 점에 주목할 필요가 있다.

**빅데이터의 다양성**

빅데이터에서 다양성은 조금 다른 의미가 있다. 생태계에서 종의 다양성이 생태계를 건강하게 하는 것처럼 데이터의 다양성은 데이터가 만들어 내는 정보의 가치를 건강하게 만든다. 우리가 선거에서 여론조사를 할 때 표본의 다양성에 대해서 이야기하는 것과 같은 이치다.

즉 데이터의 다양성이 떨어지면 한쪽에서 제공하는 일방적인 데이터를 반영할 확률이 높다. 또한 데이터의 출처가 한정되어 있다는 것은 정형화하고 데이터를 가공, 저장, 분석하는 것이 쉽기 때문에 빅데이터보다는 전통적인 데이터베이스 기술을 적용하는 게 더 효과적이고 좋은 결과를 얻게 된다.

 **빅데이터의 변화속도**
빅데이터는 잘 가공된 데이터가 아닌 가공되지 않은 원시 데이터에서 가치를 찾는 것이다. 만약 데이터가 변하지 않는 정보라면 다양한 분석 방법으로 수차례 분석을 시도했을 것이기 때문에 새로운 정보를 얻어내기는 쉽지 않다. 하지만 데이터가 계속 변할 때에는 변화에 따라 새로운 분석 방법을 이용하고 새로운 가치를 부여할 수 있기 때문에 빅데이터 분석의 좋은 예라고 할 수 있다. 이는 빅데이터에서 항상 소셜 네트워크 서비스가 등장하는 배경이기도 하다.

소셜 네트워크 서비스는 사람과 사람의 소통 통로로서 끊임없이 새로운 데이터를 생성하고 소비하는 곳이기 때문이다. 또한 데이터의 변화 속도는 시간에 대한 반영이기도 하다. 즉 데이터가 빠르게 변한다는 것은 과거가 아닌 현재의 데이터를 담을 확률이 그만큼 더 크다는 것이다.

## 빅데이터의 규모

통계에서는 표본 데이터가 많아야 정확도가 올라간다. 분석하는 데이터의 크기가 일정 수준 이상이어야 의미 있는 데이터를 얻을 수 있다. 그러면 데이터의 크기는 얼마나 커야 할까. 데이터의 크기는 데이터가 가지고 있는 가치에 따라 다르다고 할 수 있다. 즉 찾고자 하는 특징(분석 결과)에 들어맞는 데이터가 얼마나 되는가에 따라서 다르다. 일반적으로는 100테라바이트 이상의 데이터를 빅데이터라고 이야기한다.