## 정형 데이터(Structured Data)

정형 데이터는 구조화된 데이터, 즉 미리 정해진 구조에 따라 저장된 데이터다. 표 안에서 행과 열에 의해 지정된 각 칸에 데이터를 저장하는 엑셀의 스프레드 시트, 관계 데이터베이스의 테이블이 대표적인 예시이다.

**반정형 데이터(Semi-Structured Data)**

반정형 데이터는 구조에 따라 저장된 데이터지만 정형 데이터와 달리 데이터 내용 안에 구조에 대한 설명이 함께 존재한다. 그렇기 때문에 데이터 내용에 대한 설명인 구조를 파악하는 파싱 과정이 필요하고, 보통 파일 형태로 저장된다. 반정형 데이터의 예시로 웹에서 데이터를 교환하기 위해 작성하는 HTML, XML, JSON 문서나 웹 로그, 센서 데이터 등이 있다.

**비정형 데이터(Unstructured Data)**

비정형 데이터는 정해진 구조가 없이 저장된 데이터다. 소셜 데이터의 텍스트, 이미지, 영상, 워드나 PDF 문서와 같은 멀티미디어 데이터가 대표적인 예다. SNS 이용률이 크게 높아지면서 실시간으로 많은 양의 비정형 데이터가 생산되고 있다.