

제1장 데이터의 개념 및 환경 변화

1. 데이터의 개념

1.1 데이터의 사전적 정의

- ① 이론을 세우는 데 기초가 되는 사실 또는 바탕이 되는 자료
- ② 관찰이나 실험, 조사로 얻은 사실이나 정보. '자료'의 순화
- ③ 컴퓨터가 처리할 수 있는 문자, 숫자, 소리, 그림 따위의 형태로 된 정보

1.2 데이터의 어원

DARE → DATUM → DATA
(give) (Things given) (Things given)

1.3 역사 속의 데이터

- 울산 반구대 암각화 / 빗살무늬 토기 / 파피루스: 문자의 발명

2. 데이터 환경의 변화

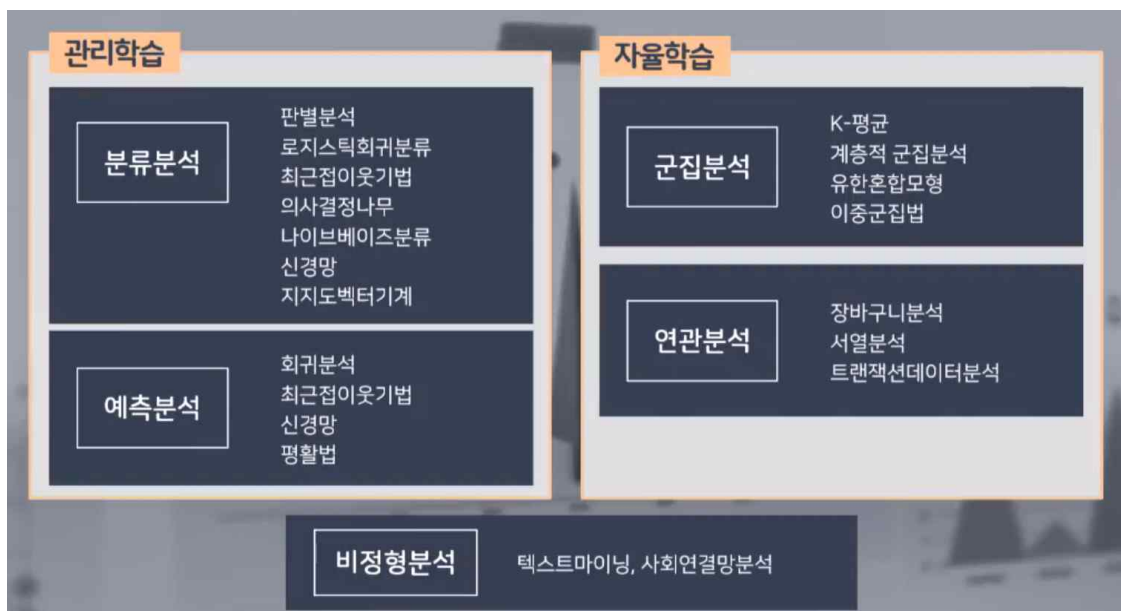
2.1 빅데이터의 출현

빅데이터의 요소: Volume(양), Variety(다양성), Velocity(속도) + Veracity(정합성)

2.2 비정형 데이터

- 비정형 데이터: unstructured data
 - 페이스북, 트위터와 같은 소셜 미디어나 웹 공간의 블로그나 게시판, 또는 인스타그램이나 유튜브와 같은 매체를 통한 사진, 동영상 데이터 등

< 데이터마이닝 기법의 분류 >



3. 데이터 활용을 위한 제도적 장치(데이터 3법)

3.1 개인정보 보호법 : 데이터 기반의 신산업 육성

3.2 정보통신망법 : 감독의 효율성 향상

3.3 신용정보법 : 개인정보 보호법과 유사·중복되는 조항 정비, 빅데이터 분석과 이용의 법적 근거를 명확화

< 데이터 3법의 주요 개정 내용 >

개인정보 보호법	<ul style="list-style-type: none">• 가명정보 개념 도입 및 활용• 개인 식별이 되지 않도록 조치를 취한 데이터를 상업적 목적으로 활용 가능• 행안부, 방통위, 금융위 등 분산된 개인정보 관리감독 기능을 개인정보보호위원회로 일원화• 정보보호 규정 위반 시 과징금 부과 등 책임성 강화
정보통신망법	<ul style="list-style-type: none">• 온라인상 개인정보 보호 관련 규제와 감독 기능을 개인정보보호 위원회로 이관
신용정보법	<ul style="list-style-type: none">• 가명정보의 금융 분야 빅데이터 분석 및 이용 가능• 가명정보를 신용정보 주체의 동의 없이 이용 가능

4. 데이터 활용의 유의점

산업과 경제에 미치는 긍정적인 면이 있다면, 미처 인지하지 못한 위험성도 존재