

★ 1.6 빅데이터 구현기술

단계	역할	활용 기술	
수집	■ 내·외부 데이터 연동 ■ 내·외부 데이터 통합	Crawling, FTP, Open API RSS, Log Aggregation DB Aggregation, Streaming	
적재	■ 대용량/실시간 데이터 처리 ■ 분산 파일 시스템 저장	Distributed File, No-SQL Memory Cached Message Queue	전처리
\bigcirc			
처리	■ 데이터 선택, 변환, 통합, 축소 ■ 데이터 워크플로 및 자동화	Structured Processing Unstructured Processing Workflow, Scheduler	
$\overline{\Box}$			
탐색	■ 대화형 데이터 질의■ 탐색적 Ad-Hoc 분석	SQL Like Distributed Programming Exploration Visualization	후처리
$\overline{\Box}$			7/15
분석	■ 빅데이터 마트 구성 ■ 통계 분석, 고급 분석	Data Mining Machine Learning Analysis Visualization	
\bigcirc			
88	■ 보고서 및 시각화 ■ 분석 정보 제공	Data Export/Import Reporting Business Visualization	활용

그림 1,15 빅데이터 아키텍처의 레이어 및 역할



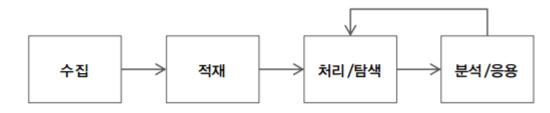


그림 1.16 빅데이터 구축 단계



lpha $oldsymbol{1}$ $oldsymbol{1}$ $oldsymbol{1}$ $oldsymbol{1}$ $oldsymbol{1}$ $oldsymbol{1}$ $oldsymbol{1}$ $oldsymbol{1}$ $oldsymbol{1}$

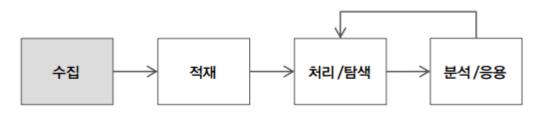


그림 1,17 빅데이터 구축 단계 - 수집 기술

표 1,3 6V 관점의 빅데이터 수집 기술

6V	수집 기술	중요성
Volume	대용량 데이터(테라바이트 이상) 수집 대규모 메시지(1,000TPS 이상) 수집	상
Variety	정형/반정형/비정형 데이터 수집 예) Log, RSS, XML, 파일, DB, HTML, 음성, 사진, 동영상 등	상
Velocity	실시간 스트림 데이터 수집	상
Veracity	N/A	하
Visualization	N/A	하
Value	N/A	하



pprox $oldsymbol{1.6}$ 빅데이터 구현기술 - 적재

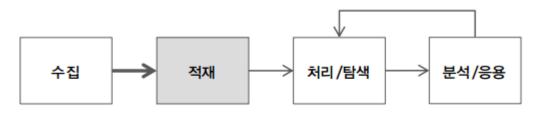


그림 1,18 빅데이터 구축 단계 - 적재 기술

표 1.4 6V 관점의 빅데이터 적재 기술

6V	적재 기술	중요성
Volume	대용량 데이터(테라바이트 이상) 적재	상
	대규모 메시지(1,000TPS 이상) 적재	
Variety	정형/반정형/비정형 데이터 수집	중
Velocity	실시간 스트림 데이터 적재	상
Veracity	데이터의 품질과 신뢰성을 확보해 적재	상
Visualization	N/A	하
Value	N/A	하



pprox 1.6 빅데이터 구현기술 - 처리/탐색

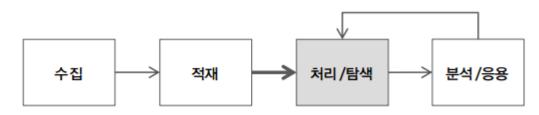


그림 1.19 빅데이터 구축 단계 - 처리/탐색 기술

표 1.5 6V 관점의 빅데이터 처리/탐색 기술

6V	처리/탐색 기술	중요성
Volume	대용량 데이터(테라바이트 이상)에 대한 후처리 및 탐색	상
Variety	N/A	하
Velocity	N/A	하
Veracity	데이터의 품질과 신뢰성을 확보하기 위한 후처리 및 탐색	상
Visualization	후처리된 데이터셋을 시각화해서 탐색	상
Value	N/A	중



★ 1.6 빅데이터 구현기술 - 분석/응용

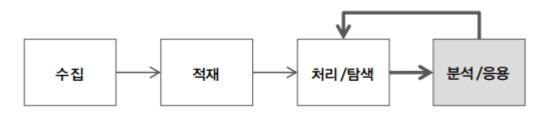


그림 1,20 빅데이터 구축 단계 - 분석/응용 기술

표 1.6 6V 관점의 빅데이터 분석/응용 기술

6V	분석/응용 기술	중요성
Volume	대용량 데이터(테라바이트 이상) 분석	상
Variety	정형/반정형/비정형 등의 다양한 데이터 분석	상
Velocity	인메모리 기반으로 실시간 데이터 분석	상
Veracity	신뢰도 높은 분석 결과를 비즈니스에 적용	상
Visualization	분석 결과 및 창출된 가치를 시각화	상
Value	분석된 결과를 비즈니스에 적용해 가치 창출	상