

165DAYS
Big Data Analysis

BIG DATA 빅데이터로 일하는 기술

감이 아닌 과학으로 성공하는 빅데이터
구축 및 활용 가이드 #청년인재 김영진

#

빅데이터가 뭐하는 것이냐고 물어볼 때마다 대답하기 힘들다. 사실 나도 잘 모른다. 통계를 공부하고 빅데이터를 접한 지 3년이 다 되어가지만 빅데이터를 정의하기가 어렵다. 예전에 여자친구가 빅데이터가 대체 뭐냐고 물어본 적이 있다.

나는 '빅데이터는 말그대로 큰 자료를 바탕으로 기업의 의사결정에 도움이 될 수

있도록 하는 기술'라고 답을 하고싶었지만 그때 당시에 나의 대답은 정말 형편

없을 정도로 간단했다. '빅데이터는 빅데이터야~' 빅데이터를 공부하고 있지

만 빅데이터에 대해 정의 내리라고 하면 확실하게 답을 하지 못했다. 왜냐하

면 통계를 배우면서 처음 들어본 단어이면서 프로그램으로 데이터를 돌려봐도

어느 정도를 빅(BIG)이라고 봐야하는지도 몰랐고, 또한 빅데이터라고 해도 이것

이 정말 가치 있는 분야인지 피부에 와 닿지 않았기 때문이다. 생각해보면 정

말로 빅데이터가 무엇인가? 라는 질문에 대한 고찰을 스스로 해보거나 정보를 찾

아본 적이 없었다. 그저 프로그램으로 데이터를 다루는게 재미있었기에 나는 유

망한 분야라는 빅데이터를 선택했을 뿐 좀 더 세세하게 역사나 탄생 배경 등

빅데이터에 대한 책 하나라도 찾아서 본 적이 없었다.

B

그런 나에게 이번에 읽은 이 책은 빅데이터에 대한 정보가 담긴
첫 빅데이터 교과서인 셈이다.

R?

책은 IT트렌드의 변화에 대해 대형 컴퓨터부터 시작해서
개인용 컴퓨터, 빅데이터 기술의 시초인 하둡(Hadoop)
의 탄생배경, 스마트폰의 모빌리티 시대와 센서의 시대까
지 발전 순서대로 서술하였다. 또한 디지털라이징 비즈니스
를 강조했는데, 나는 디지털라이징 비즈니스라는 말을 이 책을 통해 처음 알게 되었다.

디지털라이징이 무엇인가 하면, 현재의 5가지 기술(소셜, 모바일, 빅데이터, 클라우드, 사물 인터넷)을
도입해서 비즈니스를 혁신하는 것이다. 여기서 말하는 디지털라이징이 중요한 이유는 데이터양의 증가와
증가된 데이터를 활용하여 고객의 행동을 예측하고 그에 맞는 오퍼와 액션을 취할 수 있기 때문이다.
이로 인해 자연스럽게 기업경쟁력이 강화되고, 상품의 가치 또한 증대될 것이기 때문이다. 영어로 표현
되어서 다소 어렵게 느껴질 수 있으나, 간단하게 말해서 데이터를 기반으로 한 의사결정 비즈니스라고
생각하면 될 것 같다. 그리고 2장에서는 빅데이터에 대해 나를 정의를 내려준다. '빅데이터란 기존
의 기술로 처리할 수 없는 데이터로, 이전에는 그 내용을 분석할 수 없었지만 기술이 발달함에 따라 내
용 분석이 가능...' 너무 길다. 한마디로 '분석을 통한 가치 있는 데이터들의 집합'이라고 생각한다. 또
한 흥미로우면서 중요한 글귀를 보았는데 '데이터란 과거의 기록이기 때문에 후행성을 지녔지만 동시
에 미래에 일어날 일을 내포하는 선행성을 가지고 있다'는 글귀였다. 미래를 예측한다는 행위는 영
화나 만화에서만 가능한 것이 아니라, 데이터를 바탕으로 충분히 가능한 만큼 그 가치 또한 굉장하다고
할 수 있다. 빅데이터를 배우길 잘 했다고 느낀 대목이었다. `Git add . Git commit -m "First"`

165 DAYS

Big Data Analysis

책은 IT트렌드의 변화에 대해 대형 컴퓨터부터 시작해서
개인용 컴퓨터, 빅데이터 기술의 시초인 하둡(Hadoop)
의 탄생배경, 스마트폰의 모빌리티 시대와 센서의 시대까
지 발전 순서대로 서술하였다. 또한 **디지털라이징 비즈니스**
를 강조했는데, 나는 디지털라이징 비즈니스라는 말을 이 책을 통해 처음 알게 되었다.

디지털라이징이 무엇인가 하면, 현재의 5가지 기술(소셜, 모바일, 빅데이터, 클라우드, 사물 인터넷)을
도입해서 비즈니스를 혁신하는 것이다. 여기서 말하는 디지털라이징이 중요한 이유는 데이터 양의 증가
와 증가된 데이터를 활용하여 고객의 행동을 예측하고 그에 맞는 오퍼와 액션을 취할 수 있기 때문이
다. 이로 인해 자연스럽게 기업경쟁력이 강화되고, 상품의 가치 또한 증대될 것이기 때문이다. 영어로
표현되어서 다소 어렵게 느껴질 수 있으나, 간단하게 말해서 데이터를 기반으로 한 의사결정 비즈니스라
고 생각하면 될 것 같다. 그리고 2장에서는 빅데이터에 대해 나를 정의를 내려준다. '빅데이터란 기
존의 기술로 처리할 수 없는 데이터로, 이전에는 그 내용을 분석할 수 없었지만 기술이 발달함에 따라
내용 분석이 가능...' 너무 길다. 한마디로 '**분석 가능하며 가치 있는 세상 모든 정보들의 집합**'이다. 또
한 흥미로우면서 중요한 글귀를 보았는데 '데이터란 과거의 기록이기 때문에 후행성을 지녔지만 동시
에 미래에 일어날 일을 내포하는 선행성을 가지고 있다'는 글귀였다. 미래를 예측한다는 행위는 영
화나 만화에서만 가능한 것이 아니라, 데이터를 바탕으로 충분히 가능한 만큼 그 가치 또한 굉장하다고
할 수 있다. 빅데이터를 배우길 잘 했다고 느낀 대목이었다.

165 DAYS

Big Data Analysis

