



# 전문가 리포트 01

## 데이터 경제 시대 미디어 비즈니스 전략

### 1. 들어가며

### 2. 미디어 산업과 데이터

- 2.1. 전 세계 미디어산업 개관
- 2.2. 미디어 산업과 빅데이터

### 3. 데이터와 미디어 비즈니스 전략

- 3.1. 미디어 산업에서 데이터의 잠재력
- 3.2. 목표에 따른 데이터 전략

### 4. 마치며

# 데이터 경제 시대 미디어 비즈니스 전략

조대근(서강대학교 공공정책대학원 겸임교수)

## 요약문

미디어의 디지털화, 미디어 소비의 개인화, 소비자 단말기의 첨단화는 미디어 부분에서 데이터 활용을 가능하게 한 요인일 뿐 아니라 미디어 사업자들에게는 필수적인 전략으로 자리 잡았다. 미디어 사업자의 사업 목적에 적합한 데이터 활용을 위해서는 데이터 가치사슬별 중점적으로 접근하는 전략적 접근이 필요하다. 본고는 빅데이터가 가진 4가지 특징(4V), 미디어 기업들의 5가지 공통된 사업 목적, 5단계의 데이터 가치 사슬을 고려한 전략적 접근방안을 소개하고 있다. 이번 정부에서 발표한 ‘디지털 미디어 생태계 발전방안’을 통해 데이터 활용을 제약하는 요인(개인정보 유출, 저작권 침해, 전문가 부족 등)들이 최소화되고 그 결과 미디어 사업자들의 전략적 빅데이터 활용 활성화로 이어지길 기대한다. 이는 이용자들의 편익 증대, 미디어 기업들의 성장, 건강한 미디어 생태계 형성에 기여할 것이기 때문이다.

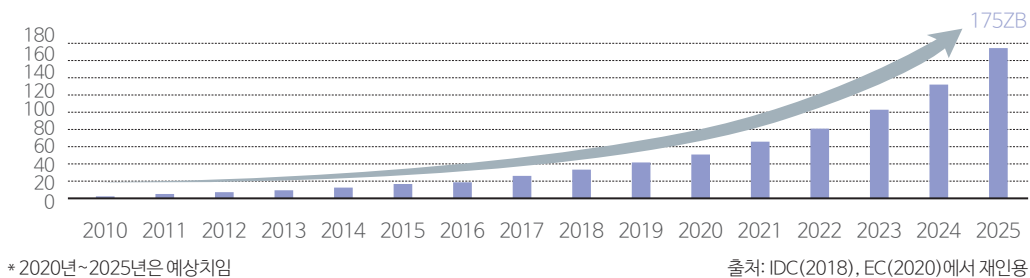


## 1. 들어가며

데이터 경제(Data Economy)<sup>1</sup> 시대이다. 현시대를 데이터 경제로 명명할 수 있는 것은 거의 모든 산업분야의 생산요소로서 데이터가 투입된다는 공통점 때문이다. 즉 데이터 경제란 디지털 기술을 이용하여 데이터를 생성, 수집, 저장, 처리, 유통함으로써 데이터를 수익 창출(Data Monetization)을 위한 수단으로 사용하고, 데이터가 자본(Data Capital) 또는 자산(Data Asset)으로 여겨지는 경제 생태계를 말한다. 그리고 데이터로 혁신을 꾀해내고 새로운 가치를 창출해 내는 것을 특징으로 한다.

데이터 경제를 실감할 수 있는 것은 데이터 트래픽의 양이라 할 수 있다. 유럽집행위원회(European Commission, 이하 EC)는 IDC 보고서를 인용하며 전 세계 생산되는 데이터가 2018년 33ZB(zettabytes)에서 2025년 175ZB까지 증가할 것이라고 하며 완전한 데이터 경제 시대에 진입했다고 보고 있다.<sup>2</sup>

그림 1 전 세계 연간 데이터 증가량 추이(FY 2010~2025년)



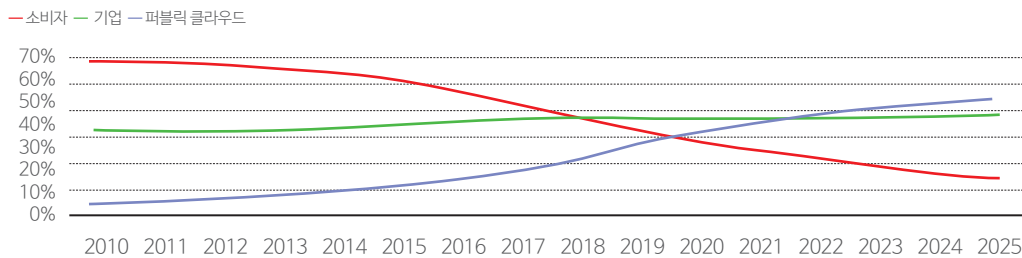
1) 데이터 경제에 대해서는 다양한 정의가 존재한다. (한국정보화진흥원, 2018) 기존의 인터넷경제, 디지털경제라는 표현보다 데이터 경제라는 표현을 즐겨하는 유럽집행위원회(European Commission, 이하 EC)는 데이터 경제를 데이터, 데이터 기술, 데이터 제품 및 서비스에 기반을 둔 경제라고 정의하고 있고, 유럽의 데이터경제는 유럽단일시장(Single Market) 역내에서 데이터 및 데이터를 통해 가능한 재화 및 서비스 이용, 수요를 증대시키는 것을 목적으로 한다고 설명한다. (EC, 2020)

2) 데이터양 단위에 관해서는 다음을 참조. 1 ZB = 103 1 EB = 106 1 PB = 109 1 TB = 1012 1 GB = 1015 1 MB = 1018 1 KB = 1021 bytes.

데이터 처리와 분석의 흐름도 변하고 있다. 데이터 처리와 분석의 80%는 데이터센터 및 중앙컴퓨팅설비를 통해 이루어지는 반면, 20%가 스마트기기(자동차, 가전, 제조 로봇, 최종이용자 인근 컴퓨팅 설비(엣지컴퓨팅)를 이용할 것으로 추정된다.(EC, 2020) 특히 퍼블릭 클라우드(public cloud)<sup>3</sup> 이용이 지속적으로 증가하여 2025년에는 전체 데이터 이용량의 약 49%를 차지할 것으로 예측된다. 지역적으로는 중국과 아시아태평양 지역의 성장이 두드러져 2025년에는 중국과 아시아태평양(중국 제외) 지역의 데이터 규모 비중이 각각 30%, 20% 인접한 수준을 보이지만 미국은 20% 미만으로 지속적인 하락세를 보일 것으로 추정된다.

그림 2 전 세계 데이터 저장 장소 이용 이용량(FY 2010~2025년)

(단위: %)

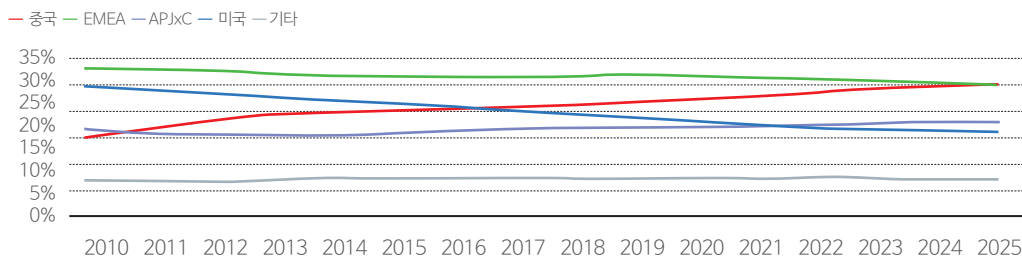


\* 2020년~2025년은 예상치임

출처: IDC(2018)

그림 3 전 세계 지역별 데이터 생성 규모 추이(FY 2010~2025년)

(단위: %)

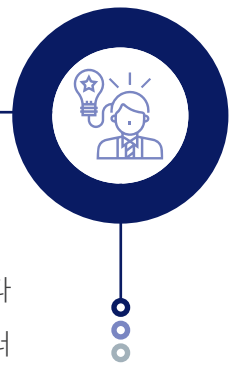


\* 2020년~2025년은 예상치임

\*\* EMEA: Europe, the Middle East and Africa(유럽, 중동, 아프리카), APJx: 아시아태평양(일본 포함, 중국 제외)

출처: IDC(2018)

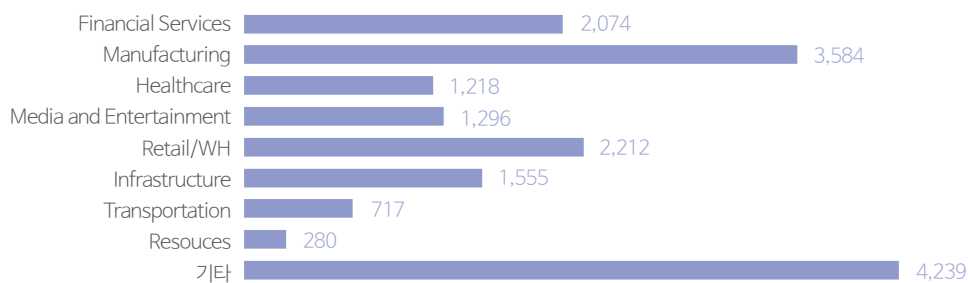
3) 클라우드 서비스는 적용하는 범위에 따라 퍼블릭(Public), 프라이빗(Private)과 이 두 가지의 혼합된 형태인 하이브리드(Hybrid)가 있다. 이 중 퍼블릭 클라우드 서비스는 일반 사용자에게 공개되어 대규모로 이루어지는 클라우드 서비스로 포털 사이트처럼 외부 데이터센터를 이용하는 유틸리티 서비스 형태로서 누구나 이용가능하다. 따라서 활용도가 높고 최소 투자로 최대 효과를 거둘 수 있는 반면, 이용자의 통제 권한이 부족하고 전문적 서비스 제공이 어렵다는 단점이 있다.(박태훈, 박경혜, (2011) 퍼블릭 클라우드 서비스의 비즈니스 모델 분석, 한국경영정보학회 학술대회, 233~239.)



데이터 규모 급증은 각 산업에도 큰 영향을 미치고 있다. 개별 산업별로 보면 그 영향의 정도는 다르지만, 영향을 받지 않는 산업이 거의 없을 정도이다. IDC가 수행한 4개 산업부문에서의 데이터 영향 분석 결과(소위 'DATCON index'를 말한다.)를 보면 제조업, 도/소매, 금융, 미디어/엔터테인먼트 순으로 데이터 규모가 큰 것으로 조사되었다. 산업별 데이터 스피어의 데이터 성장 속도(CAGR)를 보면 헬스케어가 36%, 제조 30%, 금융, 26%, 미디어/엔터테인먼트가 25%로 전반적으로 매우 급속히 증가할 것으로 예측된다.

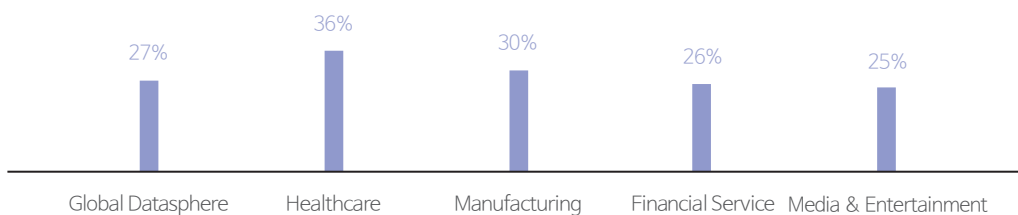
IDC의 조사 결과에서 보듯이 미디어/엔터테인먼트 산업은 타 산업 못지않게 데이터 스피어 성장 순위가 높고, 데이터에 영향을 많이 받는다. 따라서 미디어/엔터테인먼트 산업은 데이터 경제시대에 생존을 넘어 성장하기 위해서는 소극적인 대응을 넘어 적극적인 활용하는 것이 불가피하다고 할 수 있다.

**그림 4 2018년 산업별 데이터 스피어(Datasphere)<sup>4</sup> 규모**



출처: IDC(2018)

**그림 5 2018~2025년 산업별 데이터 스피어(Datasphere) 성장률(CAGR)**



출처: IDC(2018)

4) 일련의 디지털 콘텐츠를 생성, 캡처, 복사하는 등의 과정에서 발생하는 모든 데이터를 데이터스피어(Datasphere)라고 한다.

주지하듯이 미디어/엔터테인먼트 산업의 가장 주목되는 트렌드 중 하나가 “개인화(personalization)”이다. 즉 사업자가 제공하는 콘텐츠를 소비할 때 소비자들은 자기만의 소비방식을 선호하며, 앞으로는 이것이 더욱 심화될 것으로 예상된다. 미디어/엔터테인먼트 사업자들은 개인화 트렌드를 선도하지 못한다면 향후 시장에서 소비자들에게 선택받지 못할 수 있고 이는 결국 생존을 담보할 수 없는 상황에 봉착할 것이다.

그렇다면 어떻게 해야 하는가? 다수의 전문가들과 연구 결과들은 이구동성으로 그 해답은 소비자 더 정확히는 소비자로부터 얻는 데이터에 있다고 답한다. 미디어 사업자는 고객으로부터 수집한 데이터를 통해 고객을 이해할 수 있고, 매출을 증대할 수 있으며, 업무 프로세스를 개선할 수 있고, 보다 합당한 의사결정을 할 수 있다.

이에 본 글에서는 데이터의 기본 특징 4V를 개관하고 이러한 특징에 기반을 두어 미디어 사업자들이 지향하는 목표에 따라 데이터의 가치사슬(Value Chain)에 따라 전략적으로 중점을 두어야 하는 영역을 제시함으로써 보다 효율적으로 효과적으로 데이터를 활용할 수 있는 방안을 모색해 보고자 한다. 제2장에서는 미디어/엔터테인먼트 산업의 현황과 데이터의 4대 특징을 살펴본다. 제3장에서는 데이터 특성을 고려하여 데이터 가치사슬별 미디어/엔터테인먼트의 데이터 활용 전략을 정리한다.



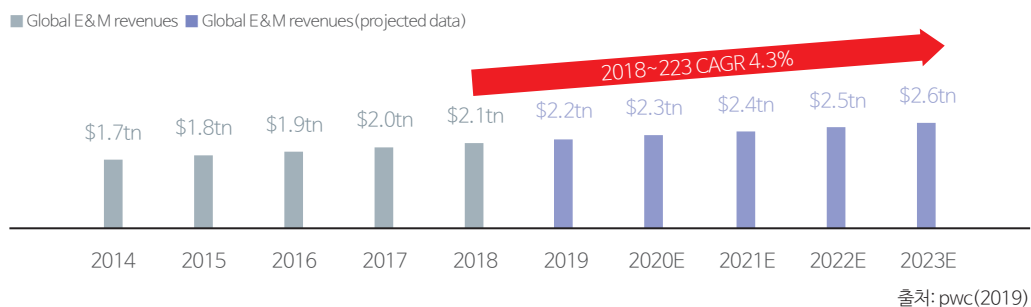
## 2. 미디어 산업과 데이터

### 2.1. 전 세계 미디어산업 개관

미디어 산업에서 주목할 점은 첫째로 시장규모 확대가 지속될 것으로 예상된다. 전 세계 미디어/엔터테인먼트 시장 규모는 2.3조 달러이고 이는 2023년 2.6조 달러까지 성장할 것으로 예측되고 있다. 2023년까지 전 세계 미디어/엔터테인먼트 산업은 연평균(CAGR) 4.3%로 전반적인 경제 성장률보다 높을 것으로 보인다. (pwc, 2019) 여기에 더하여 최근 코로나 19로 인한 비대면(Untact) 문화가 확산되면서 인터넷 기반의 미디어 콘텐츠 소비가 더욱 증가하며 미디어 산업의 지속적인 성장이 가속화될 것으로 보인다.<sup>5</sup>

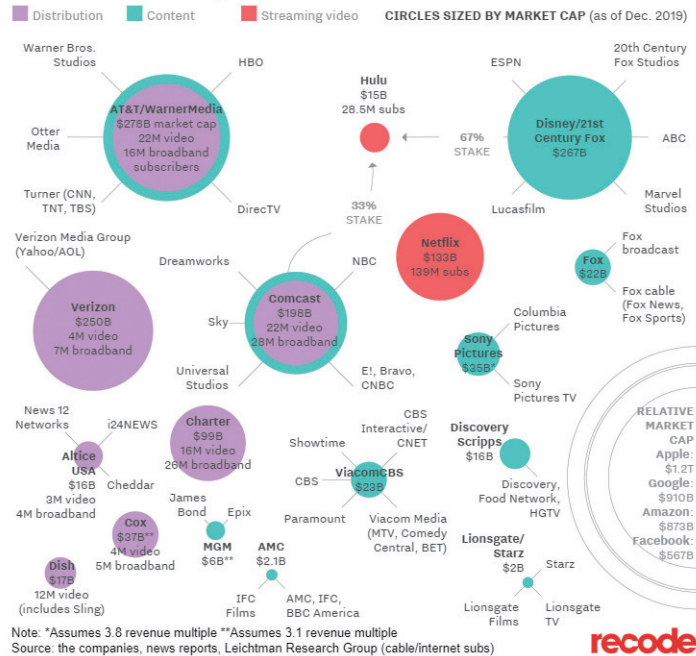
둘째, 공급 측면에서 전 세계의 주요 미디어 기업은 이전부터 전략적 M&A를 통해 규모를 확대하고 시장 내 선점, 경쟁력 확대를 도모하고 있다. Disney가 폭스('17.6월), Hulu('19.5월) 인수를, 통신사업자인 AT&T가 Time Warner(HBO 등)를 인수('19년)하였다. 그리고 Netflix는 '13년 24억 달러에서 '19년 150억 달러로 투자 규모를 확대하였다. 또 애플은 '19년 60억 달러 투자한 바 있는데 이 모두는 콘텐츠 확보를 위한 매우 공격적인 투자의 전형이다.

**그림 6 전 세계 미디어/엔터테인먼트 산업 규모 추이(FY 2014~2023년)**



5) 사실이 같은 지점은 2020.6.22. 발표된 정부의 “디지털 미디어 생태계 발전방안”에서 언급된 사항이기도 하지만 현실에서도 확인되고 있다. 예를 들면, 2020년 1분기 Netflix는 약 1,600만 명에 가까운 신규 가입자가 증가하였다. 이는 2020년 1월, 코로나 19 위기가 닥치기 전 2019년 4사분기 순증 가입자 876만 명을 고려할 때 두 배 이상 많은 수치다. 비슷한 시기 NewYork Times가 미국의 인터넷 이용을 분석한 그래프를 보면 Netflix(16%), 유튜브(15.3%), 페이스북(27%)의 트래픽이 급증하였음을 알 수 있다. 또한 화상회의 앱인 Zoom, 화상수업을 위한 Google Class 이용도 급증한 것으로 확인되었다.

그림 7 전 세계 미디어 기업 미디어 지형(2019.12. 기준)



출처: Recode(2019)<sup>6</sup>

마지막으로 미디어/엔터테인먼트 수요 측면에서 핵심 트렌드는 “개인화”와 “디지털화”라고 할 수 있다.(pwc, 2019) 소비자는 자신의 디지털 단말기로 자신의 취향에 맞는 콘텐츠를 시간과 공간의 제약 없이 받고 소비하고 있다. 이 때문에 사업자들은 고객의 소비 패턴과 같은 데이터를 활용하여 개별 소비자의 선호를 반영한 콘텐츠를 만들어 소비자 전체가 아닌 각각의 개별 고객에게 접근하고 있다. AI 알고리즘에 의한 자동추천<sup>7</sup>, 개인화된 마케팅 기법 적용을 통한 매출 증대 활동을 꾀하는 이유도 이 때문이다. 이 모든 것은 미디어/엔터테인먼트 산업 내 디지털화 진전 그리고 그로 인해 수집 가능해진 데이터로 인해 가능해진 것이다.

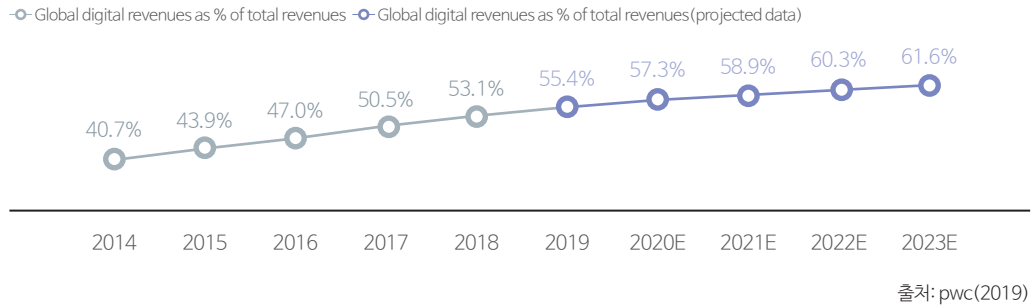
6) Rani Molla and Peter Kafka(2019), Here's who owns everything in Big Media today, VOX, 2019. 12.5.

7) 대표적인 예가 Netflix 사례다. 한 보고서의 평가 문구를 보면 이를 확인할 수 있다. “Netflix의 성공신화는 콘텐츠추천서비스에서 비롯되었다 라고 해도 과언이 아니다. 최근 Netflix에서 대여되는 영화의 75%는 머신러닝에 의한 추천으로부터 발생하며 개인화된 비디오 추천은 신규 비디오 혹은 전체 평점이 높은 비디오 추천보다 2~3배의 시청효과가 있는 것으로 나타났다.”(삼정KPMG 경제연구원, 2018)





**그림 8 미디어의 디지털화 추이(FY 2014~2023)**



## 2.2. 미디어 산업과 빅데이터

이와 같은 미디어/엔터테인먼트 산업 흐름을 생각해 보면 사업자들의 데이터의 전략적 활용은 선택이 아닌 필수가 되었다. 즉 이용자에게 디지털콘텐츠를 제공, 중개함으로써 수익을 창출하는 미디어 산업은 이용자로부터 막대한 규모의 데이터(소위 빅데이터)를 수집하고 이를 가공, 처리하여 다양한 목적에 활용해야만 하는 것이다.

미디어 사업자들이 공통적인 고민은 갈수록 경쟁이 치열해지는 환경에서 운영비용을 절감하는 동시에 다양한 플랫폼과 제품을 통해 콘텐츠와 데이터를 제공함으로써 수익을 창출해야 한다는 것이다. 이를 위해서는 데이터가 가진 특징을 잘 이해하고 목적에 맞는 전략적 활용이 필요하다.

### ① 빅데이터의 특징: 4V

이를 위해서는 데이터(최근에는 이를 빅데이터로 표현)의 특징을 먼저 이해할 필요가 있다. 통상 빅데이터의 특징을 4V로 표현하는 데 데이터양(Volume), 속도(Velocity), 다양성(Variety), 가치(Value)를 말한다.(표 1 참조)<sup>8</sup>

8) 2001년 Laney가 Gartner 블로그에 올린 글에서 빅데이터의 개념 요소로서 크기(volume), 속도(velocity) 및 다양성(variety) 등 '3V'를 제시하였는데, 그 후 Stucke와 Grunes가 가치(value)를 추가하여 '4V'를 제시하였다. (OECD, 2016)

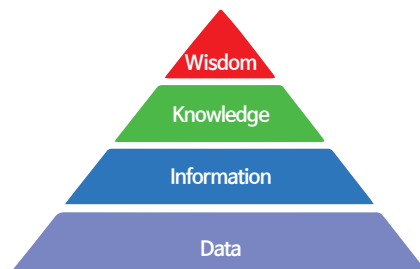
표 1 빅데이터의 특징: 4V

빅데이터 특징	주요 내용
Volume(양)	데이터의 규모가 매우 커서 TB, PB, ZB 등으로 표기될 정도의 규모. 데이터 볼륨의 증가는 무어의 법칙으로 가능해짐. 즉 기술발전으로 전 세계 거의 모든 개인(소비자)이 접근할 수 있는 더 강력하고, 더 작고, 더 똑똑하고, 더 저렴한 장치를 사용하면서 생산되는 데이터 규모는 막대해짐
Velocity(속도)	데이터 처리속도가 거의 실시간에 육박. 데이터의 실시간 처리는 새로운 가치 창출의 동인으로 작용
Variety(다양성)	데이터의 다양성 또한 수집과 처리 능력에 힘입어 증가하여, 기업은 고객의 주소(물리적 위치 또는 IP), 생일, 성별뿐만 아니라, 가구 구성, 식습관, 구매 이력, 물리적 방문 빈도 및 지속 기간과 같은 여러 가지 다른 정보 확보 가능. 더 나아가 소비자의 선호 등 고객 프로필 확보가 가능해짐
Value(가치)	데이터 자체는 수집 방법에 따라 '무료'로 간주될 수 있지만, 데이터에서 정보를 추출하는 프로세스는 가치를 창출한다는 의미

출처: OECD(2016)

Martha L. Stone(2014)은 “빅데이터는 미디어 기업들에는 큰 기회지만, 일반적으로 미디어 기업들은 여전히 빅데이터 전략의 최대 수혜를 추출하는 데는 아직 멀었다. 빅데이터 전략이 매출로 계층구조 피라미드와 같이 세분화한다면 미디어 기업의 수준은 정보(information)와 지식(knowledge) 단계 중간 정도 수준에 불과하다. 지혜(wisdom)의 단계에 도달하기 위해서는 미디어 기업들은 빅데이터 전략의 기회를 수용하고, 투자를 활용하기 위해 기술과 교육에 투자하며, 마지막으로 이 새로운 지혜를 고객 서비스나 상품과 같은 프로세스를 개선하는 데 사용할 필요가 있다.”고 진단한 바 있다.

그림 9 매슬로 계층구조로 나타난 빅데이터 단계



출처: WNMN and Freeworld Media(2014), Martha L. Stone(2014)에서 재인용



## ② 데이터원(源)과 빅데이터 특징

Helen Lippell(2016)에 따르면 미디어 부문에서 생성되는 데이터는 크게 미디어 사업자의 서비스를 직접 이용하는 이용자로부터 확보하는 “내부생성데이터(Internally generated data)”와 간접적으로 획득하는 “제3자 데이터(Third-party data)”로 구분된다. 데이터 출처와 빅데이터 4대 특징을 연결해 보면 미디어부문에서 데이터를 어떻게 이용해야 하는지에 대한 실마리를 모색할 수 있다.

**표 2 빅데이터의 4대 특징과 연결한 미디어 데이터원(源)**

내부생성데이터	핵심 “V” 특징(Key “V” characteristic)
이용자 서비스 상호작용을 포함한 이용자 프로필	<ul style="list-style-type: none"> <li>• Volume: 대량의 데이터가 저장되고 마이닝할 수 있음.</li> <li>• 이용자가 미디어제공사업자와 상호작용하는 방식에 따라 상이한 방식을 고려할 때는 다양성(Variety)이 적용됨.</li> <li>• 사업자들이 수집한 정보를 조합(join up the dots)하면 이용자를 보다 잘 이해할 수 있는 기회 도출</li> </ul>
네트워크 로깅(Network logging, 예: 웹이나 엔터테인먼트 기업이 자체 네트워크 운용)	<ul style="list-style-type: none"> <li>• 속도(Velocity): 네트워크 이슈는 문제 해결과 고객 신뢰 유지를 위해 실시간 확인</li> </ul>
기관이 최종이용자에게 제공되는 데이터 서비스 보유	<ul style="list-style-type: none"> <li>• 특징은 데이터에 대한 사업 목적에 따라 달라짐. 예를 들면, 뉴스 기관은 이용자에게 빨리 전달하는 것이 가장 우선(속도, Velocity)이며, 방송사는 복수의 단말기 유형에 복수의 포맷으로 전달되는 스트리밍 콘텐츠에 집중할 것(다양성, Variety)</li> </ul>
클릭 스트림 데이터, 상품 사용 행태, 구매이력 등을 포함하여 데이터원(source)에서 추론되는 소비자 선호	<ul style="list-style-type: none"> <li>• Volume: 대량의 데이터 수집 가능</li> <li>• Velocity는 서비스가 사용자 행동에 대응해야 하는 경우에 적절(예: 플레이어에게 추가 기능을 상향 판매하는 온라인 게임 네트워크)</li> </ul>
제3자 데이터	<ul style="list-style-type: none"> <li>• 핵심 “V” 특징(Key “V” characteristic)</li> </ul>
상업데이터 제공(Commercial data feeds), 스포츠 데이터, 언론에이전시 뉴스와이어	<ul style="list-style-type: none"> <li>• 속도(velocity): 스포츠나 뉴스 이벤트와 같은 데이터를 가장 먼저 사용하는 것이 경쟁 우위 구축하는 것</li> </ul>
네트워크 정보(외부 네트워크 사용, 예를 들면, 모바일 네트워크 상의 메시지 앱)	<ul style="list-style-type: none"> <li>• 속도(velocity): 서비스 연속성을 위해 네트워크 이슈는 실시간 확인 필요</li> </ul>
공공부문 데이터 개방	<ul style="list-style-type: none"> <li>• 진실성(veracity): 공개 데이터에는 품질, 입증(출처), 완전성 문제가 있을 수 있음</li> </ul>
자유 구성 그리고/또는 연결된 데이터. 예를 들면 위키피디아/DBpedia	<ul style="list-style-type: none"> <li>• 진실성(veracity): 클라우드 소싱된 데이터에는 품질, 입증, 완전성 문제 처리 필요</li> </ul>
소셜미디어 데이터. 예를 들면, 업데이트, 동영상, 이미지, 링크, “좋아요”와 같은 시그널	<ul style="list-style-type: none"> <li>• 양, 다양성, 속도, 진실성(Volume, variety, velocity, and veracity): 미디어 기업들은 예상되는 이용 사례에 따라 처리의 우선순위를 정해야 함.</li> <li>• 한 예로 데이터 저널리즘은 분석과 해석을 위해 대량의 데이터 준비 필요</li> <li>• 반면에, 미디어 마케팅 사업은 다수 채널에 걸친 다양한 소셜 데이터에 더 관심을 가질 수 있음</li> </ul>

출처: Helen Lippell(2016)

## 3. 데이터와 미디어 비즈니스 전략

### 3.1. 미디어 산업에서 데이터의 잠재력

미디어 부문이 빅데이터를 활용할 때 기대할 수 있는 잠재력은 크게 3가지로 알려져 있다. 이 잠재력들은 미디어 사업자들이 빅데이터를 전략적으로 이용해야 하는 이유이기도 하다.

- 첫째, 제품 및 서비스 제공 능력 향상이다. 즉 빅데이터를 활용하는 미디어 사업자들은 보다 정교한 방식으로 콘텐츠를 큐레이션 할 수 있는 역량을 갖추 가능성이 높다.
- 둘째, 소비자와 사업자 간 긴밀한 관계 형성에 기여한다. 즉 미디어 기업들이 빅데이터를 활용하여 고객에 대한 선호도, 프로필, 태도 등 더 많은 정보를 얻고, 그 정보를 활용하여 보다 긴밀한 관계를 구축할 수 있다는 것이다. 인터넷에 액세스하는 대부분의 이용자들이 소셜 미디어와 데이터 캡처 도구를 이용할 수 있는 상황에서 이들은 콘텐츠 소비자이면서 미디어 사업자에게 콘텐츠를 공급하는 역할을 하고 있어 양자 간의 상호작용은 더욱 긴밀해 지고 있다.
- 셋째, 인프라 및 프로세스상의 개선이다. 신생기업과 중소기업은 오픈소스와 클라우드 인프라를 이용함으로써 빅데이터 시대에 대응하지만, 빅데이터 활용을 위해 전통적 IT 인프라를 보유하고 있는 규모가 크고 오래된 기업들은 프로세스와 조직 문화 개선을 도모할 수 있다. 미디어 부문이 데이터 주도 사업모델을 지향함에 있어 적응하지 못했을 때 기업의 규모가 크다고 하여도 도태될 수 있기 때문에 그 잠재성을 활용해야 하는 영역이기도 하다.



### 3.2. 목표에 따른 데이터 전략

#### ① 빅데이터 활용 목적과 가치사슬

미디어 부문 내 사업모델이 다양하지만 각 사업모델을 운영하는 기업들의 목적은 몇 가지로 수렴할 수 있다. 특히 앞서 제시한 빅데이터의 잠재력을 고려할 때 미디어 부문 플레이어들의 공통된 목적에 따른 전략을 검토해보는 것이 유용하다. 즉 미디어 부문에서 빅데이터를 활용하는 5가지 목적을 달성하기 위해 빅데이터의 가치사슬 중 어디에 집중해야 하는지를(특히 기술적 측면에서) 도출하는 것이다.<sup>9</sup> 5가지 목적이란 비즈니스 프로세스 개선, 의사결정 개선, 매출 증대, 비용절감, 고객 이해를 말하고, 빅데이터의 가치사슬은 데이터 획득, 분석, 큐레이션, 저장, 이용 등이다. 우선 빅데이터의 가치사슬에 대한 의미를 살펴보면 표 3과 같다,

**표 3 빅데이터의 가치사슬**

가치사슬	주요 내용
데이터 획득	<ul style="list-style-type: none"> <li>데이터 수집은 데이터 웨어하우스나 데이터 분석을 수행할 수 있는 다른 스토리지 솔루션에 데이터를 넣기 전에 데이터를 수집, 필터링 및 정리하는 과정</li> </ul>
데이터 분석	<ul style="list-style-type: none"> <li>데이터 분석은 획득한 원시 데이터를 도메인별 사용하는 것 외에 의사결정에 사용할 수 있도록 하는 것과 관련</li> <li>데이터 분석은 관련 데이터를 부각시키고, 비즈니스 관점에서 잠재력이 높은 유용한 숨겨진 정보를 합성, 추출하는 것을 목표로 하여 데이터를 탐색, 변환, 모델링하는 것을 포함(관련 분야로는 데이터 마이닝, 비즈니스 인텔리전스, 머신러닝 등이 있음)</li> </ul>
데이터 큐레이션	<ul style="list-style-type: none"> <li>데이터 큐레이션은 효과적인 사용에 필요한 데이터 품질 요구 사항을 충족하기 위해 수명 주기 동안 데이터를 적극적으로 관리하는 것(Pennock 2007).</li> <li>데이터 큐레이션 프로세스는 콘텐츠 생성, 선택, 분류, 변환, 검증, 보존 등 다양한 활동으로 분류 가능</li> <li>데이터 큐레이션은 데이터의 접근성 및 품질을 개선하는 전문 큐레이터에 의해 수행</li> <li>데이터 큐레이터는 데이터가 신뢰할 수 있고, 검색 가능하고, 접근 가능하고, 재사용 가능하고, 목적에 맞는지 확인할 책임이 있음</li> </ul>
데이터 저장	<ul style="list-style-type: none"> <li>데이터 스토리지는 데이터에 대한 빠른 액세스가 필요한 애플리케이션의 요구사항을 충족하는 확장 가능한 방식으로 데이터를 지속성 및 관리하는 기능</li> </ul>
데이터 이용	<ul style="list-style-type: none"> <li>데이터 사용에는 데이터에 대한 액세스가 필요한 데이터 중심 비즈니스 활동, 데이터 분석 및 비즈니스 활동 내에서 데이터 분석을 통합하는 데 필요한 툴이 포함</li> <li>기업의 의사결정에서 데이터를 이용함으로써 비용 절감, 부가가치 증대 또는 기존 성과 기준에 대해 측정할 수 있는 다른 매개변수를 통해 경쟁력 강화</li> </ul>

출처: Edward Curry(2016)

9) 이하 내용은 Helen Lippell (2016)이 The Media and Entertainment Sectorial Forum에 참여한 전문가들과 토론을 통해 도출한 것을 정리한 것이다.

## ② 빅데이터 활용 목적별 중점을 뒀야 하는 데이터 가치사슬

미디어 부문 기업들의 데이터 활용에 따른 5가지 목적별로 데이터 가치사슬 중 추진해야 하는 기술적 전략 요건을 정리하면 다음과 같다.

**표 4 사업프로세스 개선 (Improve business process)**

빅데이터 요건	획득 (Acquisition)	분석 (Analysis)	큐레이션 (Curation)	저장 (Storage)	이용 (Usage)
콘텐츠와 오리진에 무관한 방식으로 이질적인 데이터 소스를 큐레이션	×		×		
트렌드에 대해 프로그램화된 방식으로 데이터 수집		×			
새로운 데이터 유형이 필요할 때 신속하게 처리	×	×	×	×	
텍스트의 정서, 토픽, 기타 무형적 측면과 관련된 비정형 데이터 분석		×	×		
형식, 의미 및 품질과 관련 공공부문의 개방형 데이터 전환 및 확대	×	×	×	×	
검색 애플리케이션을 위한 확장 가능한 툴				×	×
애널리틱, 메트릭 데이터 시각화 (특히 비즈니스 기술 사용자의 경우)		×			×
메타데이터를 자동으로 생성하고 데이터셋에 적용	×	×	×		

**표 5 의사결정 개선 (Improve decision-making)**

빅데이터 요건	획득 (Acquisition)	분석 (Analysis)	큐레이션 (Curation)	저장 (Storage)	이용 (Usage)
거의 실시간으로 빠르고 정확하게 데이터 처리	×	×		×	
데이터에 모델 및 온톨로지 <sup>10</sup> 를 적용하여 관계 추출	×	×		×	
센서에서 엑서너블 뷰로 스트림 변환	×			×	
비 프로그래머 또는 통계학자의 강력한 쿼리 및 조작을 지원하는 애널리틱 도구		×		×	
시맨틱 그래프 데이터를 분석하기 위한 추론 엔진		×	×		

10) 온톨로지(Ontology)란 존재하는 사물과 사물 간의 관계 및 여러 개념을 컴퓨터가 처리할 수 있는 형태로 표현하는 것. 어떤 일정 범위에서 사용되는 단어들의 개념, 특성, 연관 관계 등을 표현하여 단어에 대한 일반적 지식이 명시적으로 드러나고, 단어 간 관계 정의를 통해 문장의 의미를 파악할 수 있다. 스탠포드 대학의 그루버(Thomas A. Gruber) 교수는 온톨로지를 '일반적으로 개념화된 것을 형식적으로 명백하게 기술하는 명세(an explicit formal specification of a shared conceptualization)'라고 정의하였다(1992). 온톨로지는 인공지능(AI), 시맨틱 웹(semantic web), 자연어 처리(NLP: Natural Language Processing), 문헌정보학 등 여러 분야에서 지식 처리, 공유, 재사용 등에 활용된다. ([http://terms.tta.or.kr/dictionary/dictionaryView.do?word\\_seq=050917-4](http://terms.tta.or.kr/dictionary/dictionaryView.do?word_seq=050917-4))

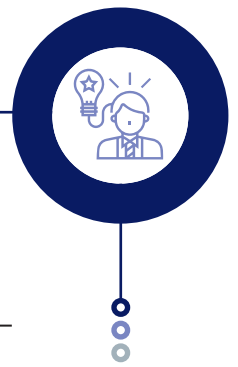


표 6 매출 증대

빅데이터 요건	획득 (Acquisition)	분석 (Analysis)	큐레이션 (Curation)	저장 (Storage)	이용 (Usage)
독점 데이터셋에서 가치 창출	×	×	×	×	×
공개 오픈 데이터셋에서 가치 창출		×	×		×
고객에게 맞춤형 데이터 및 콘텐츠 제공			×		×
인간 중심의 데이터 스트림 편집		×	×		×
데이터를 분석하여 “동일 이상의” 권장 사항을 제공하는 알고리즘					×
비기술 이용자를 위한 알고리즘 관리 도구			×		×
시맨틱 메타데이터를 사용하여 이미지 및 비디오와 같은 멀티미디어 콘텐츠 강화		×	×		×
이용자가 생성한 콘텐츠와 상업적으로 생산된 미디어의 혼합으로 새로운 디지털 제품 제작	×		×		×
데이터에서 통찰력을 생성하여 새로운 비즈니스 모델 지원 (예: 보기 습관에 따른 교차 판매)		×			×
온라인 데이터를 분석하여 오프라인 마케팅 활동 (예: 다이렉트 메일)을 통한 전환 증가		×			×
명시적으로 모델링되지 않은 트렌드, 세그먼트 및 패턴을 식별할 수 있는 예측 애널리틱 솔루션		×			
시맨틱 분석을 사용하여 소비자 대면 애플리케이션에서 보다 관련성 있는 검색 결과 제시					×

표 7 비용 절감(Reduce costs)

빅데이터 요건	획득 (Acquisition)	분석 (Analysis)	큐레이션 (Curation)	저장 (Storage)	이용 (Usage)
기존 애플리케이션보다 더 빠르게 설정할 수 있는 DB 솔루션				×	
클라우드 소싱된 데이터 큐레이션을 사용하여 내부 주제별 전문 지식 보완 역량			×		
그래프 DB의 대규모 데이터 관리				×	
비정형 데이터(예: 텍스트 또는 음성)를 하나 이상의 언어로 번역	×	×		×	
대용량 데이터 스크래핑 및 크롤링 도구	×			×	

표 8 고객이해 (Understand customers)

빅데이터 요건	획득 (Acquisition)	분석 (Analysis)	큐레이션 (Curation)	저장 (Storage)	이용 (Usage)
데이터에서 패턴을 식별하여 소비자 행동에 대한 통찰력 확보	×				
다양한 요소(예: 위치, 장치, 사용자 프로필, 사용 컨텍스트) 를 고려함으로써 콘텐츠 전달에 있어 더 나은 타겟팅	×				
모든 고객 상호작용의 데이터를 연결하여 360도 보기 구성	×	×		×	
신규 단말기에서 데이터 수집(예: 웨어러블)	×				
소비자의 행동에 대해 보다 세부적으로 검토		×			×
비정형 소셜 데이터 분석을 통해 정중 및 고객과 보다 긴밀한 관계 구축					×
개인정보 사용에 대한 명확한 정책 방향 제시				×	×

## 4. 마치며

미디어의 디지털화, 미디어 소비의 개인화, 소비자 단말기의 첨단화는 미디어 사업자들에게는 사업모델, 기술 측면에서 기존과는 새로운 접근을 요구하고 있다. 미디어 사업자들의 전략적 데이터 활용을 통한 경쟁력 확보는 선택이 아닌 필수가 되었다. 본 글은 Helen Lippell(2016)의 미디어 사업자가 추구하는 목적별로 빅데이터의 가치사슬 중 집중해야 하는 영역을 제시하고 있다. 문제는 데이터 수집 및 활용이 자유로울수록 소비자에게 보다 높은 가치를 제공할 수 있는 반면, 데이터 활용을 제약하는 요소(개인정보 유출, 저작권 침해, 전문가 부족 등)도 여전히 존재한다는 점이다. 22일 발표된 ‘디지털 미디어 생태계 발전방안’에서 제시한 정책들이 데이터 활용 제한 요인들을 최소화한다면, 미디어 사업자들의 전략적 빅데이터 활용 활성화로 이어질 것이다. 이용자들의 편익 증대, 미디어 기업들의 성장, 건강한 미디어 생태계 형성을 기대해 본다.





## REFERENCE

1. David Reinsel, John Gantz, John Rydning(2018), The Digitization of the World From Edge to Core, IDC, 2018.11.
2. EC(2020), Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions A European strategy for data, Brussels, 2020.2.19. COM(2020) 66 final
3. Edward Curry(2016), The Big Data Value Chain: Definitions, Concepts, and Theoretical Approaches, In New Horizons for a Data-Driven Economy(pp. 29-37). Springer, Cham.
4. Emilio Ontiveros, Verónica López Sabater(2018), The Data Economy Wealth 4.0, Ariel and Fundación Telefónica, 2018.4.
5. Felix Richter(2020), Netflix Sees Unprecedented Growth Amid Pandemic, 2020.4.22
6. Lippell, H.(2016). Big Data in the Media and Entertainment Sectors. In New Horizons for a Data-Driven Economy(pp. 245-259). Springer, Cham.
7. Martha L. Stone(2014), Big data for Media, Reuters Institute for the Study of Journalism, University of Oxford, 2014.11.
8. OECD Competition Committee.(2016). Big data: Bringing competition policy to the digital era. DAF/COMP (2016), 14.
9. PwC(2019), PwC Global Entertainment & Media Outlook 2019-2023
10. Rani Molla & Peter Kafka(2019), Here's who owns everything in Big Media today, VOX, 2019. 12.5.
11. 과학기술정보통신부 보도자료(2020), 청년 크리에이터와 미디어 기업이 혁신 성장을 이끄는 미디어 강국 실현- 범정부 합동'디지털 미디어 생태계 발전방안'수립 -, 2020.6.22.
12. 박태훈, 박경혜.(2011). 퍼블릭 클라우드 서비스의 비즈니스 모델 분석. 한국경영정보학회 학술대회, 233-239.
13. 삼성경제연구원(2018), 비즈니스 기회창출을 위한 AI 알고리즘의 활용, Issue Monitor 제84호, 2018.6
14. 한국정보화진흥원, 데이터 경제 기반 정책 연구, 4차산업혁명위원회, 2018.12.