

2019 KCA Media Issue & Trend

01.5

## 트렌드리포트

### AI를 활용한 콘텐츠 및 이용자 반응 분석

#### 1. 들어가며

#### 2. AI를 활용한 콘텐츠 및 이용자 반응 분석

2.1. 콘텐츠 분석 : Player Spotlight

2.2. 이용자 반응 분석 : Twitter Sentiment Analysis

#### 3. 마치며

# AI를 활용한 콘텐츠 및 이용자 반응 분석

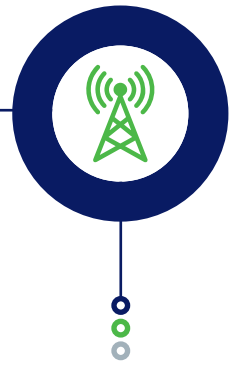
## 요약문

미디어 산업에서 AI의 활용이 확산되고 있다. 그중에서도 인간의 일상 언어를 이해하는 자연어 처리 기술의 활용도가 높다. Fox Sports의 콘텐츠 분석 서비스인 Player Spotlight와 Twitter의 이용자 반응 분석 기술인 Twitter Sentimental Analysis 등이 그 예이다. Player Spotlight는 선수 개개인의 경기 데이터와 비교 데이터 등의 정보를 기존보다 상세하고 세분화해 실시간으로 제공하고 있으며, Twitter는 인공지능 Sentiment Analysis를 활용하여 특정 인물, 제품, 서비스에 대한 이용자의 반응을 분석하고 있다.

## 1. 들어가며

방송 산업계에서 인공지능(Artificial Intelligence, 이하 AI)에 대한 관심이 날로 높아지고 있다. AI는 미디어 기업이 생산하는 콘텐츠의 질을 높이는 다양한 혜택을 제공한다. AI는 데이터에 숨겨진 의미를 추출하고, 분석 과정을 가속함으로써 미디어 콘텐츠를 풍부하게 할 수 있다. 미디어 업계에서의 AI의 활용은 콘텐츠 제작, 편집, 보안 등 다방면에 걸쳐 있으며 앞으로도 지속적으로 확대될 것으로 보인다. 본고에서는 AI의 활용 분야 중에서도 방송 콘텐츠 분석에 활용되는 사례를 살펴보고자 한다. 특히 방송 콘텐츠 분석 기술 중 자연어 처리 기술을 중점으로 해당 기술의 활용 사례를 고찰하고자 한다.





## 2. AI를 활용한 콘텐츠 및 이용자 반응 분석

### 2.1 콘텐츠 분석 : Player Spotlight

스포츠 중계방송은 AI 기술을 적극적으로 활용하는 분야 중 하나이다. 스포츠 중계는 실시간성이 중요한 영역이다. 이용자에게 실시간으로 경기 내용이나 선수에 대한 부가정보를 제공하는 것은 스포츠 중계방송 서비스의 큰 차별점이 될 수 있다. 이에 실시간으로 데이터를 추출하고 분석할 수 있는 AI 기술이 주목받고 있다. 실제로 다수의 방송·미디어 사업자들이 스포츠 중계에 AI 기술을 도입하고 있다. 미국의 스포츠 전문 방송사인 Fox Sports는 2019년 프랑스에서 개최된 제8회 FIFA 여자 월드컵 경기 중계에 AI를 활용한 Player Spotlight를 선보였다. Player Spotlight는 스포츠 경기에 대한 다양한 정보를 수집하고 분석해, 이용자에게 실시간으로 해당 정보를 제공하는 AI 기반 영상중계 플랫폼이다. 현재 Player Spotlight는 Fox Sports의 홈페이지와 모바일 애플리케이션에서 제공되고 있다.

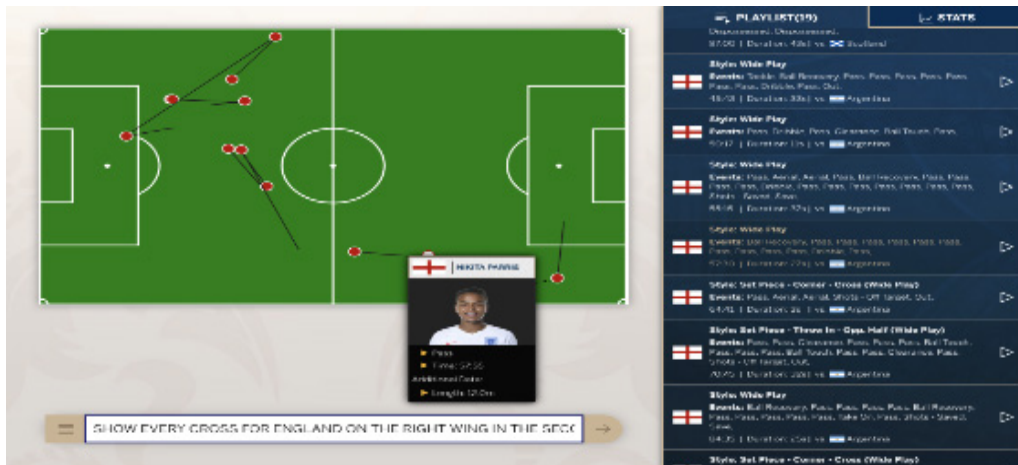
Player Spotlight 플랫폼에는 IBM이 개발한 AI 분석기술인 IBM Watson이 적용되었다. 이 기술은 선수의 움직임과 패스, 골, 킥 등 다양한 정보를 수집해 경기 데이터를 구축한다. Player Spotlight의 이용자는 타 플랫폼에서 얻기 어려운, 구체적이고 세분된 정보를 얻을 수 있다. 예를 들어 ‘2018 월드컵에서 영국 팀의 수비수가 시도한 크로스<sup>1</sup>를 모두 나타내시오’라고 IBM Watson에 입력하면 (그림 1)과 같은 결과가 나타난다.

Player Spotlight는 선수 개개인의 공 점유율 등을 추적해, 특정 선수의 경기 기여도를 예측할 수도 있다. IBM Watson에 ‘각 영국 선수의 경기 기여도를 나타내시오’ 라고 입력하면 선수별로 볼 터치, 반칙, 골, 태클 등의 기록이 나열되고 이를 종합적으로 분석해 기여도 순위를 그래프화해 보여준다.

---

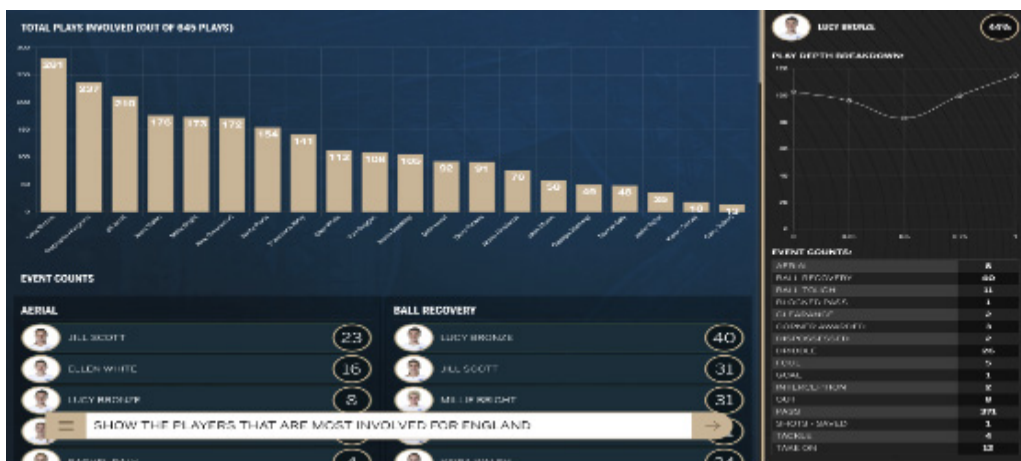
1) 필드 반대쪽이나 상대 골대에 가까운 쪽으로 공을 길게 띄워서 패스해주는 축구 기술

그림 1 IBM Watson의 경기 분석 프로세스



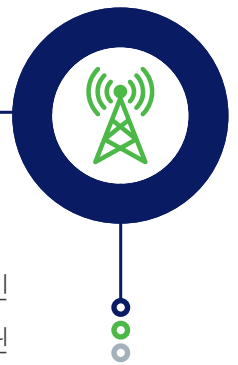
출처: IBM

그림 2 IBM Watson으로 분석한 선수 순위



출처: IBM

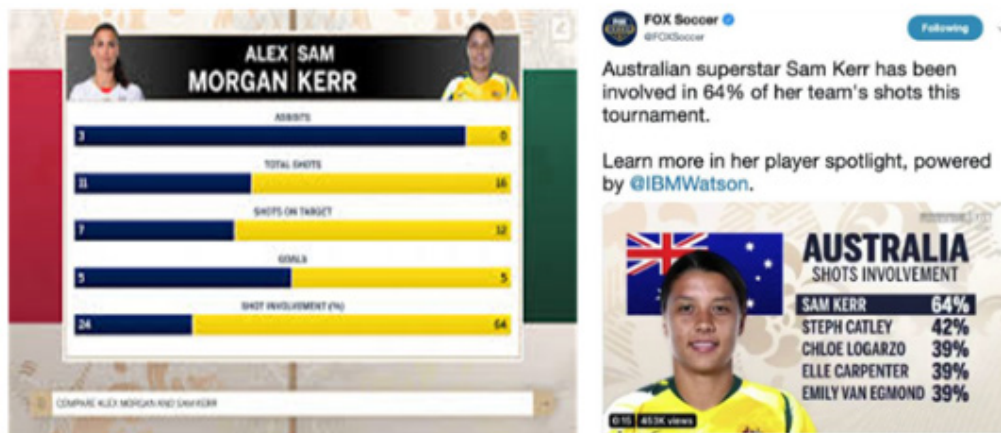
IBM Watson이 특정 데이터를 추출, 분석할 수 있는 이유는 일상 언어를 이해하고, 그 언어로 소통할 수 있는 자연어 처리 기능 덕분이다. 기계가 일상 언어를 이해하려면 우선적으로 문장의 문법 구조를 파악해야 한다. 기계는 문장에서 명사, 동사, 형용사 등의 문법적 요소를 파악해야 하며, ‘사람들’처럼 복수형으로 쓰인 문법 요소는 원형인 ‘사람’으로 재분석해야 한다. 또한, 문장의 개체명과 의도를 정확히 파악해야 한다. “오늘 서울 교통 상황은 어때?”라는 문장



을 이해하려면 1) 서울 교통 상황이 어떤지 묻는 의도, 2) 시간 개체인 '오늘'과 장소 개체인 '서울'을 모두 파악해야 한다. '미용실'과 '헤어샵'처럼 표기는 상이하지만 유사한 의미를 지닌 요소를 찾아내 동일한 표현으로 이해하는 기술도 필요하다.

IBM Watson은 방대한 언어 데이터를 분석하고 학습하는 머신러닝(machine learning)을 반복하여 일상 언어 학습 경험을 극대화한다. 위와 같이 IBM Watson이 스포츠 경기 분석에 사용될 경우에는 팀명, 선수명, 포지션명 등 경기 관련 용어 등의 머신러닝을 반복하는 것이다. IBM Watson이 분석한 정보는 실시간으로 중계 화면에 반영된다. 하지만 IBM Watson이 분석한 방대한 정보를 중계 화면 안에 모두 담기는 어렵다. 이에 Fox Sports는 자사 Twitter 계정을 통해 IBM Watson이 분석한 경기 기록을 추가로 제공하고 있다.

그림 3 Player Spotlight에서 제공한 경기 영상



출처: IBM

스포츠 경기의 실시간 정보는 기존 스포츠 중계 영상에서도 제공되어 왔지만, 각 팀의 점유율, 슈팅 개수 등 거시적 정보에 국한된다는 한계를 지녔다. 반면에, Player Spotlight는 선수 개인의 경기 데이터와 개인별 비교 분석 데이터 등 상세하고 세분화된 정보를 제공한다. 이는 자연어를 처리하는 AI 기술의 발전 덕분이다.

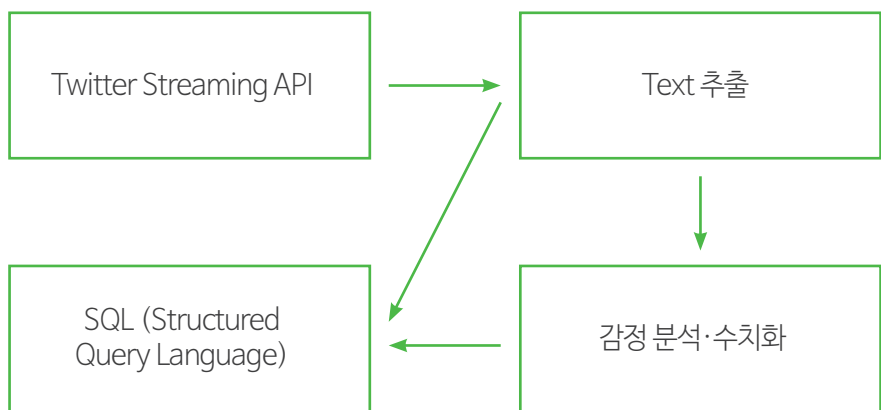
Nielsen Media Research의 발표에 따르면 2019년 7월에 FIFA 여자 월드컵 결승전을 Fox

Sports 채널로 시청한 미국인은 1,427만 명에 이르며, 실시간 인터넷 방송 시청자 또한 4년 전인 2015년 결승전보다 400% 상승했다. 시청률 상승은 OTT 확산의 영향이 크겠으나, AI 기술을 활용한 실시간 경기 정보 제공 또한 간과할 수 없는 요소이다.

## 2.2 이용자 반응 분석 : Twitter Sentiment Analysis

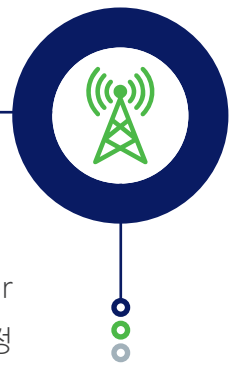
각종 사업자들이 소셜 미디어 플랫폼을 제품, 서비스 등의 광고 수단으로 활용하면서, 소셜 미디어 플랫폼이 기업과 소비자가 상호작용하는 공간으로 거듭나고 있다. 이에 소셜 미디어 플랫폼인 Twitter는 특정 인물, 제품, 서비스에 대한 이용자의 반응을 평가하는 솔루션인 Twitter Sentiment Analysis를 기업 및 일반인에게 제공하고 있다. Twitter Sentiment Analysis는 해시태그(#)가 작성된 데이터<sup>2</sup>에서 감정적 표현을 추출, 분석하여 선호도를 파악하는 콘텐츠 분석기술로, 데이터의 문맥과 상황을 종합하여 감정의 미묘한 차이를 파악해야 하기 때문에 매우 복잡한 알고리즘을 사용한다. ‘좋다’, ‘싫다’ 와 같이 직설적 표현뿐만 아니라 ‘잘도 한다’ 와 같은 반어적 표현까지도 이해해야 하기 때문이다.

그림 5 Twitter Sentimental Analysis 처리과정



출처: Medium

2) 일례로, 넷플릭스 콘텐츠 중 <Stranger Things>에 대한 Twitter 사용자 반응을 알고자 하는 경우 '#Stranger Things' 라는 해시태그가 작성된 데이터를 추출



(그림 5)는 Twitter Sentimental Analysis 처리 과정을 도식화한 것이다. Twitter Streaming API<sup>3</sup>에서 데이터를 전송받은 후, 이중 텍스트 자료를 선별해 감정 분석 과정을 거쳐 특정 숫자로 수치화한다. 수치화한 자료는 그래프로 시각화한 후 SQL<sup>4</sup>에 저장한다. Twitter Sentiment Analysis는 긍정적 감정과 부정적 감정을 서로 다른 숫자로 분류하여 수치화하는데, 숫자 분류 기준은 이용자가 자율적으로 설정할 수 있다. 이용자는 긍정적·부정적 감정을 수치화할 때 숫자 0과 1을 활용할 수 있으며, 긍정적·중립적·부정적 감정을 각각 -1, 0, 1 등으로 설정할 수 있다. 이렇게 수치화에 쓰이는 기술이 자연어 처리 기술이다. 머신러닝을 통해 데이터 표현을 학습한 Twitter Sentimental Analysis는 (그림 6)처럼 1과 0으로 수치화된 데이터를 제공하며, 해당 데이터는 그래프로 추출가능하다(그림 7).

**그림 6** Twitter Sentiment Analysis의 수치화된 데이터

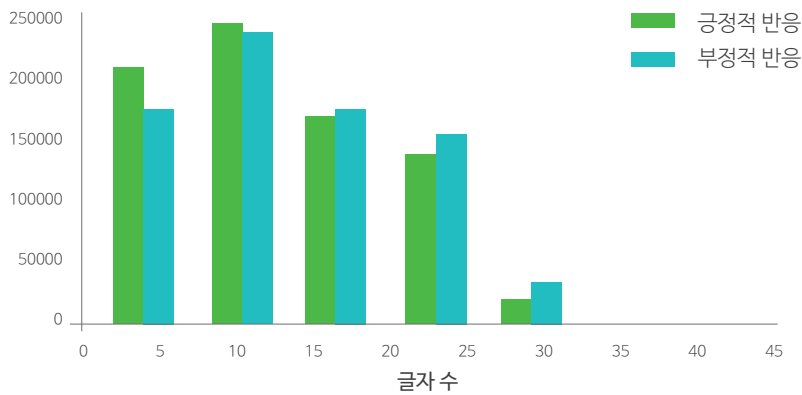
label		text
0	0	is so sad for my APL friend.....
1	0	I missed the New Moon trailer...
2	1	omg its already 7:30 : O
3	0	... Omgaga. Im sooo im gunna CRY. I've been at this dentist since 11 .. I was suposed 2 just get a crown put on (30 mins)...
4	0	i think mi bf is cheating on me!!! T_T
5	0	or i just worry too much?
6	1	jooooooooooooooooooooooooooooo Chillin!!
7	0	Sunny Again Work Tomorrow : -   TV Tonight
8	1	handed in my uniform today. i miss you already
9	1	hmmmm... i wonder how she my number @-)

출처: Towards Data Science

3) Application Programming Interface, 응용 프로그램에서 사용할 수 있도록, 운영 체제나 프로그래밍 언어가 제공하는 기능을 제어할 수 있게 만든 인터페이스

4) Structured Query Language, 데이터베이스를 추가, 삭제, 편집할 때 사용하는 명령어

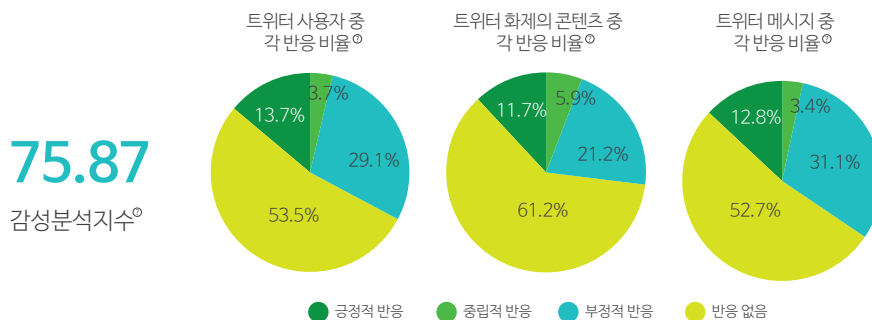
그림 7 Twitter Sentiment Analysis의 그래프화된 데이터



출처: Towards Data Science

Twitter의 월간 이용자 수는 1억 3,000만 명을 상회한다(2018년 2월 기준). 특정 제품, 서비스에 대한 Twitter 이용자 반응이 메타 데이터가 될 수 있는 규모이다. Twitter Sentimental Analysis는 특정 방송 콘텐츠에 대한 반응을 추정하는 데도 유용하게 쓰일 수 있다. 앞서 소개한 2019년 FIFA 여자 월드컵의 이용자 반응을 Twitter Sentimental Analysis로 분석한 결과, 월드컵에 대해 중립적인 입장을 표시한 콘텐츠는 52.7%로 가장 많았으며, 긍정적인 반응은 12.8%, 부정적 반응은 3.4%로 나타났다. 이처럼 Twitter Sentimental Analysis를 활용해 방영 예정인 콘텐츠에 대한 반응을 사전에 짐작할 수 있으며, 부정적인 반응이 우세한 경우에는 콘텐츠 개선안을 미리 마련하는 데 도움이 된다.

그림 8 Twitter Sentimental Analysis를 이용하여 2018년 러시아 월드컵 관련 Twitter 메시지를 분석한 결과



출처: Tweet Binder





### 3. 마치며

AI 기술은 미디어 영역에서도 다양하게 활용되고 있다. 본고에서는 AI 기술의 미디어 활용 영역 중에서도 콘텐츠와 이용자 반응 분석에 활용되는 사례를 소개했다. 특히 기계가 인간의 일상 언어를 이해하도록 돕는 기술인 자연어 처리 기술은 Fox Sports가 제공하는 Player Spotlight와 Twitter가 제공하는 Sentimental Analysis에서 유용하게 활용되고 있다.

전체적인 경기 상황·선수 개개인의 데이터를 폭넓게 제공하는 Fox Sports는 향후 스포츠 중계 기술의 판도를 바꿀 수 있다. 또한, 특정 콘텐츠에 대한 사용자 반응을 분석하는 Twitter Sentimental Analysis는 방송 콘텐츠에 대한 사전 반응과 선호도를 파악하는 데 유용하게 쓰일 수 있다. 자연어 처리 기술은 방송 시장에 적용된 AI 기술의 일부일 뿐이며, 앞으로 AI 기술이 적용된 방송·미디어 업계의 진화가 어디까지 이루어질지 기대되는 상황이다.

#### REFERENCES

1. IBM 홈페이지.  
<https://www.ibm.com/blogs/emerging-technology/ibm-player-spotlight-built-with-ibm-watson/>  
<https://www.ibm.com/blogs/emerging-technology/raise-their-game/>  
<https://www.ibm.com/blogs/emerging-technology/ibm-player-spotlight-built-with-ibm-watson/>
2. Insights Atlas (2017.06.01.) 7 Benefits of Sentiment Analysis You Can't Overlook.  
<https://blog.insightsatlas.com/7-benefits-of-sentiment-analysis-you-cant-overlook>
3. LA Times (2019.07.08.) Viewership of Women's World Cup final exceeds 14 million on Fox.  
<https://www.latimes.com/sports/soccer/la-sp-world-cup-tv-ratings-20190708-story.html>
4. SK 공식 블로그 홈페이지.  
<https://blog.skcc.com/2808>
5. Towards Data Science 홈페이지.  
<https://towardsdatascience.com/the-real-world-as-seen-on-Twitter-sentiment-analysis-part-one-5ac2d06b63fb>
6. Tweet Binder 홈페이지.  
<https://www.tweetbinder.com/blog/Twitter-sentiment-analysis/>