

# 빅데이터 분석을 활용한 복싱 관련 뉴스 게재 경향 분석

이상호\*(강남대학교 교수)

## 국문초록

본 연구는 복싱과 관련한 사회현상에 대하여 신문에서 뉴스보도가 어떻게 이루어지고 있는지를 살펴보기 위해 한국언론진흥재단의 빅카인즈 시스템을 활용하였다. 구체적으로 1990년부터 2020년 6월까지 복싱의 뉴스 게재 경향을 알아보기 위해 신문기사 빅데이터를 중심으로 신문사별, 연도별, 분야별로 분석하였고 게재 경향을 비교 및 분석하였다. 특히 언론의 뉴스보도가 현실 그 자체라기보다는 언론이 재구성한 현실이라는 전제하에서 일반 공중은 물론 사회적 인식의 형성에도 영향을 미칠 수 있는 신문기사의 언어적 상징과 의미구조를 빅데이터를 통하여 찾고자 하였다. 첫째, 약 30년 간의 키워드 ‘복싱’ 과 관련된 기사는 1,623건이며, 2008 베이징 올림픽에서의 백종섭 선수의 8강전 기권과 김정주 선수의 동메달 획득으로 인해 2008년에도 보도량이 증가한 후, 2012 런던 올림픽과 연예인인 이시영의 복싱 관련 기사로 인해 2012년에 보도량이 빠르게 증가한 것으로 나타났다. 둘째, 복싱 관련 뉴스에서 제일 큰 비중을 차지하는 주제는 ‘스포츠’ 인 것으로 나타났고, 그 뒤로 ‘사회’ 와 ‘국제’ 가 높은 비중을 차지하는 것으로 나타났다.

주제어 : 빅데이터 분석, 복싱, 뉴스 게재 경향

---

\* 이 논문은 2019학년도 강남대학교 교내연구비 지원에 의해 수행되었음.

\* run2025@gmail.net

# I. 서 론

## 1. 연구의 필요성 및 목적

한국 스포츠의 발전은 1960년 초 국민체육진흥법 조성을 시작으로 스포츠 전담 정부기구가 조직된 후 꾸준한 성장을 거듭해왔다(박진경, 2008; 체육백서, 2015). 또한 세계에서 4대 메가스포츠 이벤트를 모두 개최하게 되는 5번째 국가로 거듭나면서 스포츠 강국으로 인정받기 시작하였다(구창모, 2010; 김영갑, 2015). 그러나 한국 스포츠는 눈부신 발전과 외형적 성장에도 불구하고 국민복지, 전인교육, 국위선양 등의 전통적인 스포츠 패러다임을 기반으로 새로운 문화로써 성장하였으나 학교체육, 생활체육, 전문체육의 독립적이고 상호배타적 관계가 고착화되어 기형적이고 왜곡된 형태의 발전이 지속되고 있다(박진경, 2008).

특히 비인기 스포츠 종목의 문제는 상황이 심각하다. 대부분의 비인기 종목은 대중의 무관심, 정부의 열악한 지원과 관리, 선수수급의 어려움이 있으며, 국내 스포츠계의 균형적인 발전을 위해 빠른 시일 내에 해결되어야 할 난제로 남아있다. 대표적인 국내 비인기 스포츠 종목으로 복싱을 예로 들 수 있다. 복싱은 대한민국이 광복을 한 이후 올림픽에서 메달을 획득한 최초의 종목 중 하나였고, 1986년 서울아시안게임에서는 12체급을 모두 석권하였으며, 1988년 서울올림픽에서는 12개의 금메달 가운데 2개의 금메달을 획득하는 등 효자종목으로서의 기능을 발휘하였다(김호섭, 김진표, 김주영, 2017).

그러나 1980년대 초부터 4대 프로스포츠가 출범한 후 복싱의 인기가 점차 수그러들었고 현재는 방송국에서조차 복싱중계를 회피할 정도로 좋지 않은 처지에 놓여있다(김원기, 전만중, 박상호, 2003). 복싱단체에 등록된 선수를 보면 2011년 2,431명, 2013년 2,305명이었고, 2015년에는 2,077명으로 나타나면서 지속적으로 감소하고 있는 실정이다(문화체육관광부, 2015).

2015년 전 세계인의 이목이 집중된 스포츠이벤트 가운데 하나가 파퀴아오와

메이웨더의 웰터급 통합 타이틀전이었다. 당시 국내 시청자의 평균 시청률이 12.3%이며, 순간 최고 시청률이 21.3%를 기록할 정도로 큰 주목을 받았다(닐슨 코리아, 2015). 이러한 복싱이 현재에는 다양한 프로스포츠의 출현, 복싱선수의 사망사고로 인한 위험성 논란, 앓아진 선수층으로 인한 경기력 수준 저하현상, 승부조작, 대전료와 스타선수의 부재 등으로 인하여 대중들의 외면을 받게 되었다(김원기, 전만중, 박상호, 2003; 문화일보, 2016; 장한곤, 2015). 더불어 각종 종합격투기의 출현으로 선수유출과 복싱관련단체의 대립 및 갈등은 한국 복싱의 위기상황을 한층 더 고조시키고 있다(안병오, 박성현, 조춘환, 2016).

이러한 복싱의 비인기화 문제를 해결하기 위한 방안으로 빅데이터(big data) 분석은 미래예측과 마케팅을 위한 분석기법으로 활용되고 있다. 정보기술의 발전으로 인해 데이터의 양이 빠르게 증가하여 이를 저장, 처리하는 기술도 빠르게 발전하고 있다. 빅데이터는 단순히 많은 양의 데이터라는 개념이 아닌 데이터를 수집, 저장, 분석하여 현상을 이해하고 나아가 미래 예측 기술을 통칭하며, 현재 사회 전반에 활용되고 있다(한승희, 2018). 또한 초연결사회를 살고 있는 현대에 연결된 모든 것으로부터 생성된 데이터가 빠르게 증가하고 있어 이를 처리하고 해석하여 적용하기 위한 빅데이터 기술은 앞으로도 꾸준히 사회적으로 관심을 받을 것이다. 이러한 빅데이터의 발전은 다양한 연구 분야에서도 빅데이터의 처리기술을 개발하고 적용하는 연구가 꾸준히 증가하고 있다. 빅데이터를 활용한 연구는 주로 대용량 데이터의 처리를 통하여 표본 기반 연구에서 얻을 수 없었던 데이터의 숨겨진 의미를 해석하는데 의의를 둔다(이상호, 2019).

그동안 다양한 학술 분야에서 빅데이터를 활용하여 연구가 진행되어 왔으나, 최근 뉴스 빅데이터를 활용하여 특정 사회적 이슈 분석을 위한 내용분석의 연구가 학계에서 관심을 받고 있다. 뉴스 자료는 시계열로 축적된 양질의 데이터 집합이다. 과거에는 처리시간과 정교함으로 인해 소규모 뉴스 기사 컬렉션만으로 내용 분석 연구를 수행하였지만, 최근 자연어의 처리 기술과 텍스트 마이닝 등과 관련한 기술의 발전으로 인해 대량의 기사를 처리하고 분석하는 일이 한층 수월해졌다(박대민, 2016). 이러한 기술의 발전은 뉴스 빅데이터를 활용한 연구의 활성화를 가져오고 있다.

이러한 빅데이터 분석의 장점에도 불구하고 그동안 복싱과 관련된 선행연구에서는 주로 운동과학(권경배, 2006; 김상용 등, 2015; 임용규, 김성주, 2008), 사회과학(윤현수, 2014; 정찬교, 김규호, 박승한, 2011), 심리학(이충섭, 2014; 천길영, 김주영, 2009)에서 주로 연구가 이루어지고 있다. 이 외 안병오, 박성현과 조춘환(2016)은 문헌조사와 더불어 일부 패널을 대상으로 조직관리, 정책적 측면에서 연구를 진행하였다. 이에 본 연구에서는 4차 산업시대의 핵심인 빅데이터를 활용하여 복싱 관련 기사의 정보 패턴을 분석함으로써 복싱의 문제점들에 대해 언어적 상징과 의미를 탐색하고자 하였다. 따라서 본 연구는 1990년 이후 약 30년간의 국내 뉴스 빅데이터를 대상으로 복싱 분야 보도자료 분석을 통해 복싱 관련 정보에 대한 기사 패턴과 트렌드를 파악하여 문제점을 도출하고 복싱분야의 활성화를 위한 기초자료로 활용하고자 하였다.

## 2. 연구문제

본 연구에서는 첫째, 언론보도에 나타나는 복싱 관련 키워드를 수집하고, 둘째, 수집된 키워드가 국내 주요 신문에서 어떤 관련어와 연결되어 있으며, 마지막으로 연결의 양태가 주요 프레임 차원에서 어떻게 이해되는가에 대해 조사하고자 하였다. 연구를 위해 다음과 같은 연구문제들을 설정하였다.

연구문제 1: 복싱 관련 키워드의 언론보도의 빈도는 차이가 있는가?

연구문제 2: 복싱 관련 키워드의 언론보도의 내용은 차이가 있는가?

연구문제 3: 복싱 관련 키워드의 언론보도의 양태는 어떤 프레임특성을 가지는가?

# II. 연구방법

## 1. 자료수집

본 연구는 한국언론진흥재단에서 제공하는 국내 언론사의 뉴스에 대한 데

이더베이스를 기반으로 빅데이터 기술이 도입된 분석 서비스인 빅카인즈(BIG Kinds)를 통하여 복싱에 대한 언론보도 분석을 하였다. 자료수집에 활용한 언론사는 <표 1>과 같이 서울 및 경기 등 7개 지역별 54개 언론사이며, 중앙지 11개, 경제지 8개, 지역종합지 28개, 방송사 5개, 전문지 2개이다. 검색어는 ‘복싱’ 과 ‘권투’ 로 1990년 1월 1일부터 2020년 6월 10일까지 약 30년간의 데이터를 분석하였다. 이 연구에서는 오랜 기간을 대상으로 대량의 뉴스 기사를 분석하여 기존의 신문 기사를 대상으로 진행된 내용 분석 연구와 차별하고자 하였다. 이러한 연구를 통해 뉴스보도 성향의 거시적 지형을 보는 새로운 시도가 가능하다(이은별, 전진오, 백지선, 2017). 앞에서 언급한 바와 같이 검색어와 수집 기간을 결정하여 최종적으로 1,623건의 뉴스 기사를 수집과 분석 대상으로 하였다.

표 1. 빅카인즈에서 제공하는 언론사 유형

유형	언론사
중앙지(11)	한국일보, 한겨레, 조선일보, 중앙일보, 세계일보, 서울신문, 문화일보, 동아일보, 내일신문, 국민일보, 경향신문
경제지(8)	헤럴드경제, 한국경제, 파이낸셜뉴스, 아주경제, 아시아경제, 서울경제, 머니투데이, 매일경제
지역지(28)	한라일보, 충청투데이, 충청일보, 충북일보, 중부일보, 중부매일, 중도일보, 제민일보, 전북일보, 전북도민일보, 전남일보, 울산매일, 영남일보, 부산일보, 무등일보, 매일신문, 대전일보, 대구일보, 국제신문, 광주일보, 광주매일신문, 경인일보, 경상일보, 경남신문, 경남도민일보, 경기일보, 강원일보, 강원도민일보
방송사(4)	YTN, SBS, OBS, MBC, KBS
전문지(2)	전자신문, 디지털타임스

## 2. 빅데이터 분석

빅데이터(Big data)는 대규모의 다양한 데이터로부터 가치 있는 정보를 추출하는 기술을 의미한다(Gantz, 2013). 이러한 빅데이터는 데이터의 성질이 지금까지의 다른 데이터와 비교하였을 때 근본적으로 다른 새로운 형태와 내용의 복합체로 되어 있다(한신갑, 2015). 빅데이터는 인위적 규정에 의하여 만들

어지는 것이 아닌, 일상에서 자연스럽게 만들어진 현실을 그대로 반영한 데이터라는 점에서 일반적인 사회과학 데이터와 차별성을 가진다. 산업분야와 학문분야에서 빅데이터의 중요성은 갈수록 커지고 있으며, 특히 다양하고 복잡한 문제의 해결방법을 효과적으로 제시한다는 점에서 사용이 증가하고 있다. 빅데이터는 일반적으로 양적인 개념에서 집중되어 다루어지는 경향이 있으나 본질적으로는 질적 의미가 빅데이터의 본질이라 할 수 있다(김건우, 안석준, 문혜정, 2017). 이에 복싱과 관련한 이슈들을 포괄적이고 심층적으로 알아보기 위하여 빅데이터를 이용한 언어 네트워크 분석방법은 유용할 것이다.

언어 네트워크 분석은 언어로 구성된 텍스트를 분석대상으로 하여 분석하는 방법이다. 텍스트를 중심으로 한 언어 네트워크 분석의 기본적인 요소는 연결 중심성과 근접 중심성, 매개 중심성 등의 세 가지 중심성이다. 이들 중심성 분석을 통하여 전체적인 텍스트의 의미적 특성을 파악함과 동시에 노드 간의 관계 분석이 가능하다(이소정 외, 2017). 텍스트를 이용하여 네트워크 분석을 적용한 경우 언어구조를 공간적으로 표시하여 방대한 텍스트에 나타나는 주요 키워드 간의 관계들을 시각적으로 파악할 수 있다. 또한 네트워크 중심성(network centrality)이 큰 개념을 찾아 전체 텍스트가 전달하고자 하는 의도와 의미를 이해하는 데 효율적이다(박지성, 정지원, 2013). 실제로 같은 단어라도 연도나 분류별 이미지 형성과는 평가기준이 달라 미디어의 프레임에 따라서 상이한 분석결과의 예측이 가능하다.

### 3. 텍스트마이닝

텍스트마이닝은 빅데이터를 분석하는 방법으로 다양하게 나타난 텍스트 데이터를 중심으로 의사결정을 위한 유용한 정보와 텍스트의 패턴을 도출하는 과정을 의미하는 것으로 통계학, 전산언어학, 빅데이터 분석을 포함하는 분야이다. 텍스트마이닝에서는 분류(classification)와 회귀(regression), 군집(clustering)의 기술이 적극적으로 도입되고 있다(한국재정정보원, 2019). 본 연구에서는 텍스트마이닝으로 검색어와 관련된 기사를 시계열 그래프로 확인할

수 있는 키워드 트렌드와 검색어와 연관이 있는 키워드를 워드 클라우드로 나타내는 연관어 분석을 중심으로 분석을 실시하였다.

#### 4. 분석 절차 및 평가

‘워드클라우드’는 문서에서 핵심 키워드를 추출하여 문서의 특징을 직관적으로 확인하도록 시각화하는 기법이다. 먼저 특정 키워드로 뉴스를 검색한 후 정확도를 기준에서 연구자가 설정한 건수로 뉴스 내 키워드를 추출한다. 이 후 토픽랭크(Topic Rank) 알고리즘으로 키워드 간의 상관관계를 분석하여 연관도가 높은 키워드 간에 워드클라우드로 시각화하여 나타낸다. 토픽랭크(Topick Rank)는 특정 키워드와 같이 발생된 키워드의 발생빈도 및 중요도를 계산하여 연관 키워드를 추출하는 분석으로, 핵심 개념을 추출하고자 할 때 사용된다(Bougouin, Boudin, & Daille, 2013). 본 연구에서는 가중치와 키워드 빈도수를 기준으로 분석하였다.

‘관계도 분석’은 검색한 뉴스 가운데 정확도 상위 100건의 뉴스를 기준으로 개체들(장소, 키워드, 인물, 기관) 사이의 연결 관계를 네트워크 형태로 시각화하는 방식으로, 개체명 추출방식은 SSVM(Structural Support Vector Machine)을 이용하여 기계학습 알고리즘을 사용한다.

## II. 연구결과

### 1. 보도량 분석 결과

‘복싱’에 대한 보도량 추이를 분석하고자 연도별 기사 건 수를 조사한 결과 <그림 1>과 같이 나타났다. 1990년부터 시작하여 2006년까지는 복싱과 관련된 기사 건 수가 매우 적었으나, 2007년을 시작으로 최요삼 선수의 뇌출

혈로 인한 수술과 해외 타이슨 선수의 마약복용 등의 이슈로 인해 보도량이 증가하기 시작하였다. 2008 베이징 올림픽에서의 백종섭 선수의 8강전 기권과 김정주 선수의 동메달 획득으로 인해 2008년에도 보도량이 증가한 후, 2012 런던 올림픽과 연예인인 이시영의 복싱 관련 기사로 인해 2012년에 보도량이 빠르게 증가한 것으로 나타났다. 그리고 이후 2016년까지 여러 사건들로 높은 증가추세를 보이다가 이후부터 현재까지 감소추세를 보이고 있으나, 현재는 6월까지의 뉴스만을 대상으로 하여 증가의 여지가 있다. 이를 통해 올림픽과 연예인 관련 이슈가 복싱 관련 뉴스의 증가에 영향을 미치는 것으로 유추할 수 있다.

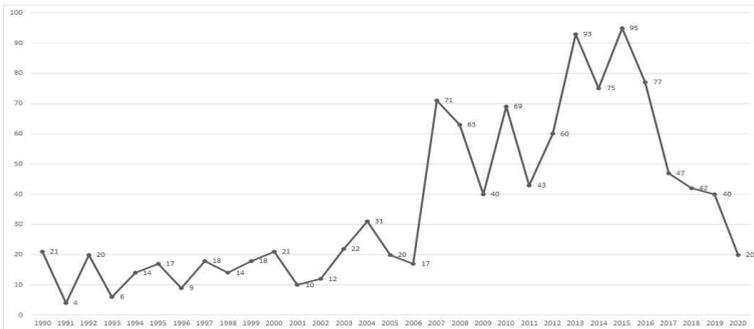


그림 1. '복싱' 관련 주제의 연간 보도량 변화 추이(1990~2020)

### 1) 주제별 보도량

‘빅카인즈’에서는 8개의 주제로 뉴스 기사를 분류하는데, 한 기사에서 3개의 주제까지는 중복 분류를 허용한다. 이 연구는 ‘빅카인즈’에서 제공하는 뉴스 분류 메타정보를 활용하여 주제별 보도량을 분석하고자 하였다. <그림 2>에서 보는 바와 같이 복싱 관련 뉴스에서 가장 많은 비중을 차지한 주제는 ‘스포츠’로 나타났으며, 그 뒤를 이어 ‘사회’와 ‘국제’가 높은 비중을 차지하는 것을 알 수 있다. ‘경제’와 ‘IT\_과학’ 주제에서는 각각 6건과 4건의 뉴스를 보도하였다.

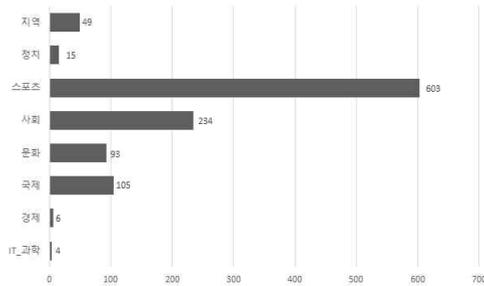


그림 2. 주제 별 '복싱'관련 뉴스 보도량

주제별 보도량의 시계열 분석을 위하여 연간 주제별 뉴스 보도량에 대한 추이를 분석하였고, 분석 결과는 <그림 3>과 같다. 전반적인 추이는 <그림 1>의 연간 보도량의 변화 추이와 비슷한 것으로 나타났다. 특히 주제 분류 비중이 높은 ‘스포츠’ 주제의 보도량 변화가 연간 보도량의 변화와 유사한 것을 알 수 있다.

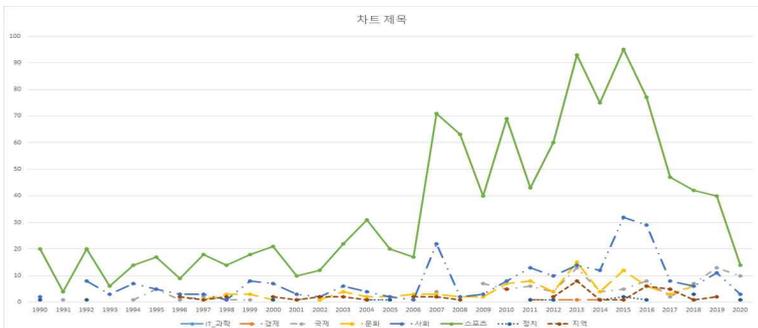


그림 3. 연간 주제별 '복싱' 관련 뉴스 보도량 변화

## 2) 언론사 유형 별 보도량

‘빅카인즈’는 언론사의 유형을 5가지로 구분하고 있다. 이 연구에서는 언론사 유형별로 기록관리 분야에 대한 뉴스 보도량을 분석하고자 하였다.

분석 결과, <그림 4>와 같이, 복싱과 관련하여 가장 많은 뉴스를 보도한 언론사는 지역종합지로 나타났고, 다음으로 중앙지인 것으로 나타났다.

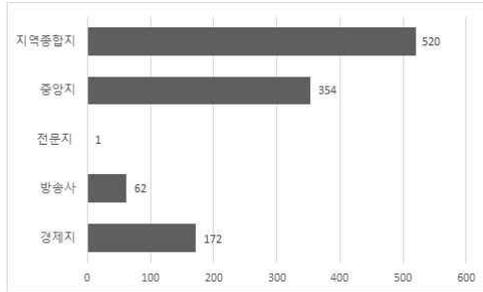


그림 4. 언론사 유형별 '복싱'관련 뉴스 보도량

그러나 <표 1>에서와 같이, 유형별 언론사의 수는 다르기 때문에 절대적인 뉴스 보도량만으로는 실제 언론사의 유형별 보도량을 알기 어렵다. 언론사의 유형별로 보도된 뉴스 기사의 건 수를 해당 유형에 대한 언론사 수로 나눈 다음 백분율로 나타내면 언론사의 유형별 뉴스 보도 점유율을 알 수 있다. 그 결과 <그림 5>와 같이, 중앙지에서 복싱 분야의 뉴스를 가장 많이 보도하였고, <그림 4>에서 가장 많은 뉴스를 보도한 것으로 나타난 지역종합지는 보도 비율이 3순위인 것으로 나타났다. 이는 지역종합지의 종류가 28종으로 가장 많아 지역종합지 1곳 당 평균 보도량이 줄어들었기 때문인 것으로 볼 수 있다.

## 2. 빈도 중심의 보도 내용 분석

연구에서 수집된 전체 뉴스를 대상으로 '빅카인즈'의 메타정보에서 제공하는 키워드를 추출하여 빈도를 계산하였으며, 이 중 상위 20위에 해당하는 단어

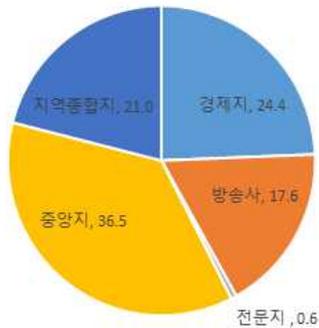


그림 5. 언론사 유형별 뉴스 보도 점유율

는 <표 2>와 같다. 빈도가 가장 높은 단어는 ‘국가대표’이며, 그 다음으로 ‘금메달’, ‘이시영’, ‘신종훈’, ‘김주희’, ‘KO’ 등의 단어가 나타났다. 상위 빈도에 나타난 단어를 통하여 유추해보면, 복싱 관련 뉴스 기사의 주요 주제가 올림픽과 관계가 있다는 것을 알 수 있다. 그리고 이는 2006년 올림픽이 시작된 이후 뉴스 기사가 많아지기 시작한 <그림 1>의 분석 결과와 일치한다.

가장 보도량이 많았던 2016년에는 2016 리우올림픽이 개최한 연도이다. 이를 반영하여 ‘국가대표’, ‘금메달’, ‘KO’ 등과 같은 단어가 높은 빈도 순위를 차지하였다. 국내 유명 복싱선수인 ‘신종훈’, ‘김주희’, ‘최요삼’, ‘함상명’ 등의 단어가 뉴스 기사에 자주 나타난 것을 확인할 수 있었으며, ‘세계챔피언’, ‘전국체전’, ‘WBC’, ‘뇌출혈’ 등과 같은 단어가 높은 빈도로 나타난 것으로 볼 때, 복싱과 관련한 일련의 이슈들이 있었으며 이들이 기사화되었음을 예상할 수 있다.

표 2. ‘복싱’ 관련 전체 뉴스 기사에 대한 상위 20위 빈도 단어 리스트

순위	단어	빈도	순위	단어	빈도
1	국가대표	236	11	WBC	61
2	금메달	179	12	동메달	59
3	이시영	156	13	판정승	56
4	신종훈	113	14	충주시청	52
5	김주희	83	15	파QUI아오	51
6	KO	80	16	인천시청	51
7	전국체전	73	17	뇌출혈	45
8	세계챔피언	70	18	은메달	38
9	최요삼	69	19	KBC	33
10	선수들	67	20	함상명	33

### 3. 복싱관련 키워드의 프레임 분석

<그림 6>은 키워드 간의 연관 관계를 분석한 후 연관도가 높은 키워드 간에 워드클라우드를 이용하여 시각화로 나타낸 결과이다.



그림 6. ‘복싱’ 관련 키워드 워드클라우드 분석 결과

가중치 기준으로 분석 시 가장 핵심적인 키워드로는 ‘금메달’, ‘이시영’, ‘국가대표’, ‘WBC’ 등이 나타났다. 이를 통해 ‘올림픽’과 ‘연예인 이시영’이 밀접하게 연관된 개념이라는 것을 확인할 수 있었다. 한편 빈도수를 기준으로 분석한 결과 핵심키워드로 ‘국가대표’, ‘금메달’, ‘이시영’, ‘신종훈’, ‘김주희’ 등 순으로 나타났다. ‘신종훈’과 ‘김주희’가 핵심 키워드로 나타난 원인은 2014년 인천아시안게임의 금메달리스트인 신종훈 선수의 태극마크 포기 사건과 김주희 선수의 세계 8대 복싱기구 챔피언 등극 등의 사건 등으로 빈도수가 높았기 때문이다.

한편 ‘복싱’ 관련 항목별 키워드의 관계도 분석결과는 <그림 7>과 같다. <인물>의 관계도 분석에서는 매개 중심성과 연관성이 높은 핵심어 중 여러 복싱 선수들과 함께 ‘전두환’이 나타난 이유는, 1986년 서울 아시안게임 당시 전두환 대통령이 메달을 딴 선수들에게 큰 액수의 포상금을 준 것에 대한 결과로 해석된다. <장소>관계도 분석에서는 대부분 복싱 대회가 열린 도시와 국가들 중심으로 분석이 되었고, <기관>관계도 분석은 매개 중심성과 연관성이 가장 높은 핵심어로 ‘검찰’이 나타났는데, 이는 2007년 대한아마복싱연맹의 훈련비 횡령 사건이 이슈화 된 결과로 보여진다. <키워드>관계도 분석에서는 ‘권투선수’, ‘대통령’, ‘올림픽’ 등의 결과가 나타났다. 이러한 인물 중심의 보도는 일화적인 프레임 속성이 강하게 내포되었다는 점을 암시한다.



## V. 결론 및 제언

본 연구는 복싱과 관련한 사회현상에 대하여 신문에서 뉴스보도가 어떻게 이루어지고 있는지를 살펴보기 위해 한국언론진흥재단의 빅카인즈 시스템을 활용하였다. 구체적으로 1990년부터 2020년 6월까지 복싱의 뉴스 게재 경향에 을 알아보기 위해 신문기사 빅데이터를 중심으로 신문사별, 연도별, 분야별로 분석하였고 게재 경향을 비교 및 분석하였다. 특히 언론의 뉴스보도가 현실 그 자체라기보다는 언론이 재구성한 현실이라는 전제하에서 일반 공중은 물론 사회적 인식의 형성에도 영향을 미칠 수 있는 신문기사의 언어적 상징과 의미구조를 빅데이터를 통하여 찾고자 하였다.

첫째, 약 30년 간의 키워드 ‘복싱’ 과 관련된 기사는 1.623건이며, 2008 베이징 올림픽에서의 백종섭 선수의 8강전 기권과 김정주 선수의 동메달 획득으로 인해 2008년에도 보도량이 증가한 후, 2012 런던 올림픽과 연예인인 이시영의 복싱 관련 기사로 인해 2012년에 보도량이 빠르게 증가한 것으로 나타났다. 이를 통해 올림픽과 연예인 관련 이슈가 복싱 관련 뉴스의 증가에 영향을 미친 것으로 예상할 수 있다. 복싱은 1970년대의 각종 스포츠의 활성화와 더불어 발전하였다가 1980년대에는 다른 스포츠와 더불어 아마복싱의 전성기를 거치면서 복싱선수가 증가하였으나, 1990년대에 들어서는 3D 현상에 따라 매맞는 운동은 기피하려는 국민정서가 크게 작용하여 그 인기가 점차 감소하였다(대한체육회, 1993). 이후 본 연구에서와 같이 올림픽을 기점으로 다시 점차 복싱의 언론노출이 많아지고 있는 것으로 나타났다.

둘째, 복싱 관련 뉴스에서 제일 큰 비중을 차지하는 주제는 ‘스포츠’ 인 것으로 나타났고, 그 뒤로 ‘사회’ 와 ‘국제’ 가 높은 비중을 차지하는 것으로 나타났다. ‘경제’ 와 ‘IT\_과학’ 주제에서는 각각 6건과 4건의 뉴스를 보도하였다. 가장 많은 양의 뉴스를 보도한 언론사는 지역종합지이며, 다음으로 중앙지가 많은 것으로 나타났다. 빈도 별 연관 단어 분석 결과에서는

‘국가대표’가 가장 높은 것으로 나타났으며, 그 다음으로 ‘금메달’, ‘이시영’, ‘신종훈’, ‘김주희’, ‘KO’ 등의 단어가 나타났다. 상위 빈도를 차지하는 단어를 통해 예상해보면, 복싱 관련 뉴스 기사의 주요 주제가 올림픽과 관계가 있다는 것을 알 수 있다. 가장 보도량이 많았던 2016년에는 2016 리우올림픽이 개최된 연도이다. 이를 반영하여 ‘국가대표’, ‘금메달’, ‘KO’ 등과 같은 단어가 높은 빈도를 차지하였다. 국내 유명 복싱선수인 ‘신종훈’, ‘김주희’, ‘최요삼’, ‘함상명’ 등의 단어 또한 뉴스 기사에 자주 나타난 것을 확인할 수 있었으며, ‘세계챔피언’, ‘전국체전’, ‘WBC’, ‘뇌출혈’ 등과 같은 단어의 빈도가 높은 것으로 볼 때, 복싱과 관련된 일련의 이슈들이 나타났고 이들이 기사화되었다는 것을 예상할 수 있다. 연관도가 높은 키워드를 워드클라우드로 시각화한 결과에서는 가중치를 기준으로 분석 시 가장 핵심적인 키워드로는 ‘금메달’, ‘이시영’, ‘국가대표’, ‘WBC’ 등이 나타났다. 이를 통해 ‘올림픽’과 ‘연예인 이시영’이 밀접하게 연관된 개념이라는 점을 확인할 수 있었다. 한편 빈도수 기준으로 분석 시 핵심키워드로 ‘국가대표’, ‘금메달’, ‘이시영’, ‘신종훈’, ‘김주희’ 등 순으로 나타났다. 이러한 결과를 통해 올림픽의 효과로 복싱 선수들이 점차 언론에 드러나면서 복싱도 부각되고 있는 것을 알 수 있다. 이에 복싱의 활성화를 위한 마케팅적 시사점을 하자면 스포츠스타의 배출로 복싱의 장점을 부각시킴으로써 점차 사회적으로 긍정적인 이미지를 구축할 수 있을 것이다.

셋째, ‘복싱’과 관련한 전체 키워드 관계도를 분석한 결과, <인물>관계도 분석에서 매개 중심성과 연관이 높은 핵심어 중 여러 복싱 선수들과 함께 ‘전두환’이 나타난 이유는, 1986년 서울 아시안게임 당시 전두환 대통령이 메달을 딴 선수들에게 큰 액수의 포상금을 준 것에 대한 결과로 해석된다. <장소>관계도 분석에서는 대부분 복싱 대회가 열린 도시와 국가들 중심으로 분석이 되었고, <기관>관계도 분석은 매개 중심성과 연관성이 가장 높은 핵심어로 ‘검찰’이 나타났는데, 이는 2007년 대한아마복싱연맹의 훈련비 횡령 사건이 이슈화 된 결과로 보여진다. <키워드>관계도 분석에서는 ‘권투선

수’, ‘대통령’, ‘올림픽’ 등의 결과가 나타났다. 이러한 인물 중심 보도는 일화적 프레임 속성이 내포되었다는 점을 암시한다.

결과적으로 복싱은 아직 예전과 같이 국민스포츠클로써 자리잡고 있지 않고 인식자체가 없거나 다양한 사건사고로 인해 부정적인 이미지로 떠올려지고 있다. 이에 복싱의 발전을 위해서는 구체화되고 정확한 복싱의 인식도 중요하다. 복싱 관련 연구를 통하여 더 많은 트렌드와 패턴을 식별하고 이를 구체적인 홍보방안으로 반영하여 다양한 문제해결을 통해 복싱의 이미지를 적극 알려야 할 것이다.

이러한 분석 결과는 언론사의 유형별 단어빈도 정보를 기반으로 내용을 분석한 연구 결과와 유사한 것으로 나타나, 언어 네트워크 분석으로 다시 한번 단어 빈도 중심의 내용 분석 결과를 입증할 수 있었다. 이 연구는 약 30년이라는 장기간의 뉴스 빅데이터를 대상으로 프레임 분석과 키워드 분석 방법론을 적용하여 최초로 복싱 분야의 뉴스 보도 특성을 중심으로 분석을 실시한 점에서 의의가 있다. 이 연구를 계기로 향후 후속 연구에서는 이 연구에서 제시하고 있는 연구 모형을 바탕으로 체육학을 구성하는 특정 주제를 중심으로 뉴스 빅데이터를 기반으로 한 분석을 시도할 수 있을 것이다.

해당 연구를 통하여 나타난 결과를 중심으로 제언을 하자면 기사의 제목을 분석하여 보다 다양하고 참신한 논의가 이루어지지 못하였다. 나아가 이 연구 결과만으로 복싱의 언론보도 특성을 일반화하기에는 어렵다는 점에서 한계가 있다. 더불어 가능한 연구자의 주관적 해석의 여지를 줄이고자 의미연결망 분석을 활용하였으나, 미디어가 재구성한 의미체계를 완전히 파악하기에는 한계가 있으며, 질적인 해석이 수반된다는 부분에서 연구자의 주관적인 해석이 완전히 배제되었다고 보기는 어렵다. 이에 향후 후속 연구에서는 뉴스가 구성한 프레임을 명확하게 추출하기 위해 기사제목 외에 기사내용과 주제, 작성자 등을 면밀히 분석하고 실제적인 공중의 인식과 비교하여 복싱에 대한 사회의 인식이 어떻게 변화되는가를 다각적으로 파악해야 할 것이다. 마지막으로 이 연구는 빅카인즈에서 검색한 빅데이터를 중심으로 분석을 하여 1990년 이전에 자료들은 검색이 불가능하여 해석에 한계가 존재하며, 일

반인과 기자의 복싱에 대한 이해 수준이 달라 해석에 대한 왜곡이 일어날 가능성이 있다.

## 참고문헌

- 구창모(2010). 비인기 종목 대학 운동선수의 생활양식 연구. **한국스포츠사회학회지**, 23(3), 1-18.
- 권경배(2006). 복싱글러브 및 헤드기어 충격량 측정시스템 개발과 충격유효성 분석. **체육과학연구**, 17(4), 1-10.
- 김건우, 안석준, 문혜정(2017). 빅데이터 분석을 통한 정치, 문화, 경제 뉴스의 미디어 프레임연구. **한국경영학회 통합학술발표논문집**, 258-281.
- 김상용, 유기웅, 김형수, 신호철, 김기형(2015). 아마추어 복싱선수의 전진스텝 유무에 따른 원·투 스트레이트 동작의 운동역학적 비교분석. **한국체육학회지**, 54(6), 587-599.
- 김영갑(2015). 한국 스포츠문화의 지향점과 선진화 대안 찾기. **한국체육학회지**, 54(1), 47-56.
- 김원기, 전만중, 박상호(2003). 한국 복싱의 퇴조에 관한 소고. **한국체육사학회지**, 8(12), 1226-4105.
- 김호섭, 김진표, 김주영(2017). 복싱 크레이닝 타격음 만족이 운동정서와 몰입에 미치는 영향. **무예연구**, 11(4), 65-84.
- 문화일보(2016). ‘한국 프로복싱 현주소…최근 10년 세계챔프 단 1명 없어…선수 대부분 생활꿈’.  
[http://www.munhwa.com/news/view.html?no=2016050401\\_033521302002](http://www.munhwa.com/news/view.html?no=2016050401_033521302002)
- 문화체육관광부(2015). **2015 체육백서**.
- 박대민(2016). 장기 시계열 내용 분석을 위한 뉴스 빅데이터 분석의 활용 가능성: 100만 건 기사의 정보원과 주제로 본 신문 26년. **한국언론학보**, 60(5), 353-407.

- 박지성, 정지원(2013). 텍스트 네트워크 분석: 사회적 인식 네트워크 (socio-cognitive network) 분석을 통한 정책이해관계자 간 공유된 의미 파악 사례. **정부학연구**, 17(1), 1-22.
- 박진경(2008). 신정부의 체육정책: 신정부 스포츠정책의 비전과 과제. **스포츠과학**, 102, 2-12.
- 안병오, 박성현, 조준환(2016). SWOT분석을 통한 대한복싱협회 경영전략. **한국엔터테인먼트산업학회논문지**, 10(6), 235-249.
- 윤현수(2014). 30대 전문직 남성의 생활체육 복싱 수련 경험에 관한 내러티브 탐구. **코칭능력개발지**, 16(2), 13-23.
- 이상호(2019). 신문의 빅데이터 분석을 통한 전통 무예의 보도 경향 및 의미연결망 분석. **무예연구**, 13(4), 233-249.
- 이소정, 김강민, 조선영, 이정수, 최송식(2017). 사회복지시설 안전관리 매뉴얼 분석-언어네트워크 분석기법을 중심으로. *Crisisonomy*, 13(5), 1-14.
- 이은별, 전진오, 백지선(2017). 서울의 다문화 공간 연구: 뉴스 빅데이터 분석 시스템 ‘빅카인즈’를 이용한 국내 언론의 외국인 마을 보도(1990~2016) 분석. **미디어 경제와 문화**, 15(2), 7-43.
- 이충섭(2014). 청소년 복싱선수들의 인지평가와 수행전략의 관계에서 정서의 매개 효과. **한국스포츠심리학회지**, 25(1), 77-88.
- 임용규, 김성주(2008). 복싱 레프트 스트레이트 동작의 운동학적 분석. **한국체육과학지**, 17(2), 1063-1072.
- 장한곤(2015). 복싱클럽 서비스품질의 중요도-만족도 평가. **한국스포츠학회지**, 13(2), 153-164.
- 정찬교, 김규호, 박승환(2011). 성인여성들의 복싱에어로빅 참여가 여가만족과 생활만족에 미치는 영향. **한국여가레크리에이션학회지**, 35(3), 79-90.
- 천길영, 김주영(2009). 복싱에어로빅 참여자의 몰입경험과 여가만족의 관계. **한국체육과학지**, 18(2), 409-420.
- 한국재정정보원(2019). **재정정보 분야 텍스트마이닝 활용 방안 연구**.
- 한승희(2018). 뉴스 빅데이터를 이용한 우리나라 언론의 기록관리 분야 보도 특성 분석: 1999-2018 뉴스를 중심으로. **정보관리학회지**, 35(3), 41-75.

- 한신갑(2015). 빅데이터와 사회과학하기: 자료기반의 변화와 분석전략의 재구상. **한국사회학**, 49(2), 161-192.
- Bougouin, A., Boudin, F., & Daille, B. (2013). Topicrank: Graph-based topic ranking for keyphrase extraction. *International Joint Conference on Natural Language Processing(IJCNLP)*, Oct 2013, Nagoya, Japan, 543-551.

## ABSTRACT

### The Trend of Posting News Related to Boxing Using Big Data Analysis

Lee, Sang-Ho(Kangnam Univ.)

This study used the Big Kinds system of the Korea Press Foundation to examine how news reports are being made in newspapers about social phenomena related to boxing. Specifically, in order to investigate the tendency of publishing news of boxing from 1990 to June 2020, it analyzed by newspaper, year, and field, focusing on big data of newspaper articles, and compared and analyzed the tendency of publishing. Especially, the study tried to find the linguistic symbol and meaning structure of newspaper articles that can affect the formation of social awareness as well as the general public through big data on the premise that news reports of the media are not the reality itself but the reality reconstructed by the media. The results of the analysis are as follows: First, the number of articles related to the keyword 'boxing' for about 30 years was 1,623, and the amount of reports increased rapidly in 2012 due to Baek Jong-seop's abstention to the quarterfinals at the 2008 Beijing Olympics and Kim Jung-ju's bronze medal acquisition. Second, the most important topic in boxing news was 'sports', followed by 'social' and 'international'. In the themes of 'economic' and 'IT\_science', 6 and 4 news reports were reported, respectively, and the most news reports were the regional comprehensive magazines, followed by the central magazines.

Key words: Big Data Analysis, Boxing, and News-Growing Trends

논문투고일 : 2020.06.30.

심사일 : 2020.07.10.

심사완료일 : 2020.07.31.