

텍스트 마이닝 기법을 활용한 태권도원에 관한 빅데이터 분석

정권혁(경희대학교 박사과정) · 전익기*(경희대학교 교수)

국문초록

본 연구는 세계 태권도인의 성지인 태권도원과 관련된 키워드를 파악하여 태권도원의 발전 방향에 관한 기초자료를 제공하는 데 그 목적이 있다. 자료수집을 위해 네이버와 다음에서 제공하는 웹, 블로그, 뉴스, 카페, 지식인, 학술정보와 구글 웹, 뉴스, 페이스북을 분석대상으로 사용하였다. 자료분석기간은 태권도원이 개장한 2014년 4월 24일부터 2019년 4월 24일까지 총 6년간의 데이터를 선정하였다. 자료수집과 분석을 위해 소셜 매트릭스 사이트인 텍스툼을 통해 키워드의 빈도와 매트릭스를 추출하고, UCINET 6 프로그램 중 Netdraw 기능을 사용하여 키워드 간의 관계와 연결중심성을 분석하고 시각화하였다. 또한, 유사한 키워드에 대한 군집을 도출하기 위하여 CONCOR분석을 실시하였다. 분석결과, 첫째, 수집된 데이터량은 총 8,283건이 도출되었다. 둘째, 태권도원의 키워드 빈도는 태권도, 무주군, 전라북도, 세계, 스포츠 순으로 나타났다. 셋째, 연결중심성은 태권도, 무주군, 전라북도, 공연, 체협 순으로 나타났다. 넷째, 키워드에 대한 N-gram 네트워크 그래프를 실시한 결과, 전라북도-무주군, 무주군-설천면, 시범-공연, 태권도-시범, 태권도-성지 순으로 나타났다. 다섯째, CONCOR분석 결과 ‘수련 및 교육장소’, ‘방문체험 콘텐츠’, ‘지자체 속성’, ‘가치속성’, ‘대회속성’으로 나타났다.

주제어 : 태권도원, 텍스트마이닝, 빅데이터

* ikjeon@khu.ac.kr

I. 서 론

태권도는 우리나라의 전통 무예이자 스포츠로써 외국에 한국 문화를 전파하는 역할을 담당하고, 강대국들과의 스포츠외교를 공유하는 역할을 수행하며, 한국을 대표하는 문화관광 콘텐츠로써 국가의 브랜드 가치에 크게 기여하고 있다(이정학, 김욱기, 2012). 정부는 ‘태권도 명품 브랜드화’, ‘태권도의 성지 설립’을 추진하였고, 이에 전라북도 무주군에 태권도 관련 체험, 교육, 연구, 국제경기, 교류의 장소를 만들고자 태권도원을 설립하였다(김종석, 정경희, 유광길, 2014).

태권도원이란 태권도를 대표하는 대형 공간으로써 국내뿐만 아니라 전 세계 태권도인의 수련과 순례의 장을 제공하고, 태권도 종주국으로써의 위상을 높이는 데 목적이 있다(김진기, 정경희, 유광길, 2017). 이에 4,500석 규모의 태권도 전용 경기장인 T1경기장을 비롯하여 400명 이상을 수용하는 실내 공연장, 1,400명의 숙박을 수용하는 태권도 연수원, 그리고 태권도로써는 세계 최대 규모인 태권도 전문 박물관, 태권도 연구소, 태권도 체험관, 본부운영센터 등이 있다(배진남, 2013). 또한, 태권도원을 관리 감독하기 위해 태권도진흥재단을 창립하여 향후 태권도원 운영과 공간계획 발전, 태권도 체험프로그램 개발 등에 관한 사업을 수행하고 있다(이미연, 전익기, 2014). 나아가 지역주민의 문화공간과 전망대 조성 등 더 많은 방문객을 모집하기 위해 노력하고 있다.

태권도원의 방문객 현황을 살펴보면, 2014년 개장한 첫해는 16만 명, 2015년에는 25만 명, 2016년에는 27만 명, 2018년도에는 32만 8천 명, 2019년도에는 31만 6천 명으로 지속적인 증가추세를 보이나 최근 2019년도에는 소폭 하락하였다(태권도진흥재단, 2020). 꾸준히 방문객 수는 증가하였으나, 한국개발원과 정부가 추정하였던 첫해 방문객 수는 178만 명으로(윤동길, 2017) 6년간의 전체 방문객 수를 합한 수치와 비슷하다. 또한, 태권도원 내 체험관 이용

객은 2014년 6,498명에서 꾸준히 증가하여 2017년 24,704명으로 정점을 찍고, 2018년 10.8% 감소한 22,047명으로 줄었다(한국문화정보원, 2019). 이렇듯 태권도원의 방문객 수가 미비한 이유는 태권도공원의 테마파크 수익 모델에 한계가 있고, 마케팅을 통한 프로모션이 미비하였으며, 태권도원의 정체성 부분에 있어서 어떠한 방향으로 진행이 되어야 하는지에 대해 명확히 구분하지 못했다는 점을 들 수 있다(김상태 외 7인, 2014). 또한, 태권도원은 실제 방문객들이 원하는 깊은 내면적 니즈를 이해하지 못하고, 단순히 외면적 확장을 통한 보여지기 식의 발전으로 인하여 정작 태권도인들이 원하는 성격의 태권도성지에 관한 의미를 제대로 파악하지 못하고 있다.

이러한 원인을 해결하기 위해 과거 선행연구에서는 태권도원의 문화관광상품 개발에 관한 연구(장혜영, 박현진, 최승희, 2016), 태권도원 문화 콘텐츠를 구축하는 방안 연구(양명모, 조성균, 2019), 태권도원 내부고객 지원활동이 시장지향성과 고객지향성에 관한 연구(김종석, 정경희, 유광길, 2014) 등 실제 태권도원이 어떠한 방식으로 가야 하는가에 관한 연구를 진행하였다. 또한, 태권도원을 방문 전과 후를 비교하여 성과 및 개선방안에 관한 연구(이미연, 전익기, 2014), 매력도 분석 통해 효율적인 전략을 탐색한 연구(유광길, 2019) 등이 진행되었으나 주로 태권도 관련 종사자들과 대학생들을 대상으로 연구하였다. 또한, 단순히 태권도원의 홍보전략(양명모, 2019; 이정희, 최문용, 2018; 장용석, 2019; 홍준기, 2016)을 연구하여 외면적 이미지를 각인시키기 위한 연구들이 진행됐다. 그러나 실제 태권도원의 내면에 내포된 의미에 관한 연구는 미비하며, 실제 방문객들이 경험한 내용의 자료는 부족한 실정이어서 이에 이용객의 방문목적 및 수요사항 반영이 필요하다.

본 연구는 태권도원에 방문한 이용객과 SNS에서 나타난 키워드를 토대로 태권도원과 관련한 빅데이터를 활용하여 자료를 수집하고, 태권도원에 대한 인식을 분석하는 연구를 진행하고자 한다. 빅데이터는 메가바이트(Mb) 단위에서 테라바이트(Tb) 단위에 달하는 거대한 데이터의 집합체로(Brown, Chui, & Manyika, 2011), 관련 단어의 정보를 추출하고, 지식 변화를 예측하며(김환표, 2013), 최근 연구에서는 소비자들의 트렌드와 인식을 통해 새로운 잠재적 변

수까지 확인할 수 있다(이정학, 이재문, 김육기, 김형근 2017).

이에 본 연구는 태권도원이 개장한 2014년부터 2019년까지 6년간 온라인에서 태권도원과 관련한 어떠한 단어들이 업로드되었는지, 실제 방문객들이 올린 후기들은 어떻게 도출되었는지를 보고, 태권도원의 마케팅 전략 도출과 발전 방향에 기초자료를 제공하는데 목적이 있다.

II. 연구방법

1. 연구대상

본 연구의 수집채널은 포털사이트인 네이버(Naver), 다음(Daum)에서 제공한 블로그, 카페, 뉴스, 웹문서, 지식인, 학술정보와 구글(Google)에서 제공한 웹문서, 뉴스, 페이스북을 선정하였다. 자료수집을 위한 단어로는 ‘태권도원’을 사용하였다. 자료 분석에 사용된 수집기간은 태권도원이 개장한 2014년 04월 24일부터 2019년 04월 24일까지 6년으로 설정하였다. 자료를 분석하기 위하여 사용된 분석 도구로는 소셜네트워크 분석프로그램인 UCINET 6과 데이터 시각화를 나타내는 Netdraw 기능을 사용하였다. 분석한 데이터의 특징은 다음 <표 1>과 같다.

표 1. 분석 데이터의 특징

구분	내용
수집채널	● 네이버(블로그, 카페, 뉴스, 웹문서, 지식인, 학술정보) ● 다음(블로그, 카페, 뉴스, 웹문서, TIP) ● 구글(웹문서, 뉴스, 페이스북)
수집기간	2014년 4월 24일~2019년 4월 24일 (총 6년)
수집도구	텍스톰(Textom)
수집단어	태권도원
단어수	키워드 빈도 상위 50개
분석도구	UCINET 6, Netdraw

2. 조사도구 및 절차

본 연구는 텍스트마이닝 기법을 사용하여 수집된 단어들의 빅데이터 분석을 통해 추출된 단어 간의 의미연결을 분석하고자 한다.

1) 텍스트마이닝(Text Mining)

텍스트마이닝은 대규모의 비정형화 데이터를 관계 데이터로 추출하는 방법으로써 텍스트에 나타나 있는 단어들을 분해하고, 특정 단어들의 출현빈도를 파악하여 단어들의 관계에 대해 조사하는 기법이다(전채남, 서일원, 2013).

본 연구에서 태권도원에 관련하여 수집된 문장 속에서 추출된 단어 중 태권도원과 무관한 단어들은 삭제하였다. 또한, 같은 의미 또는 유사하게 쓰인 단어들은 통합하였다. 마지막으로 띄어쓰기나 단어가 이루지 못한 불완전한 단어들은 정제작업을 시행하였다. 정제작업을 하는 중 연구자의 지나친 주관성을 배제하고, 신뢰성을 확보하기 위해 전문가 집단 4인(교수 1인, 박사 2인, 전문가 1인)과 함께 정제하였다. 데이터 정제는 다음 <표 2> 같다.

데이터 정제로 사용된 단어들을 살펴보면, 태권도진흥재단, 즐기다, 무주군, 열리다, 전라북도 등 같은 의미지만 명칭이 다르게 적용된 단어들은 정제하였으며, 수, 것, 옆, 작, 시, 보, 우 등 단어를 이루지 못한 것에 대해 분석이 불필요하다고 판단되어 삭제하였다.

표 2. 데이터 정제

데이터 정제	분석에 불필요한 삭제 단어
태권도진흥재단 (태권도재단, 진흥재단)	수, 것, 옆, 작, 시, 보, 우 등 의미가 우리지 못한 단어
즐기다 (즐기, 즐겁, 즐거운)	
무주군 (무주)	
열리다 (열림, 열리)	
전라북도 (전북, 전북도)	

2) 의미연결망분석(Semantic Network Analysis)

의미연결망 분석은 텍스트 내에 메시지를 형성하는 단어가 어떻게 이용되고 배치되는지에 관한 구조적 관계를 분석하는 것이며, 의미 분석 및 구조의 패턴을 통해 관계의 형태 속에서 내포된 메시지 내용을 도출하는 분석방법이다(차민경, 2015; 한관중, 2003). 이에 개인 간의 상호작용에 따라 하나의 연결망이 형성되고, 형성된 연결망 구조를 분석하여 사회적 양상이나 실태를 조사하는 방법이다(차민경, 권상희, 2015).

3. 데이터 수집

2014년 4월 23일부터 2019년 4월 23일까지 6년 동안 태권도원의 키워드를 중심으로 네이버와 다음, 구글을 통해 추출된 데이터량은 네이버 4,197건, 다음 3,517건, 구글 569건으로 수집되어 총 8,283건이 도출되었다. 정제모들은 고유명사와 복합명사를 파악하기에 적합한 Espresso K를 사용하였고, 분석품사는 명사, 형용사, 동사를 선택하였다. 자료수집 결과는 <표 4>와 같다.

표 3. 총 단어의 빈도수

구분	네이버	다음	구글
웹	1,000 (780KB)	892 (5.14MB)	387 (8.02MB)
블로그	1,000 (3.5MB)	852 (4.16MB)	-
뉴스	869 (9.57MB)	725 (5.82MB)	83 (26KB)
데이터양 카페	1,000 (424KB)	1,000 (4.77MB)	-
지식인	260 (554KB)	48 (20KB)	-
학술정보전체	68 (283KB)	-	-
페이스북	-	-	99 (46KB)
총 수집한 데이터양	4,197	3,517	569

4. 자료처리

본 연구는 텍스트(Textom)을 이용하여 포털사이트의 네이버(Naver)와 다음(Daum)이 제공한 웹문서, 블로그, 뉴스, 카페, 학술정보와 구글 웹문서, 뉴스, 페이스북에 나타난 비정형 단어를 키워드를 추출하고, 키워드의 빈도수를 통해 중요도를 파악하여, 태권도원의 세부 내용을 분석하는 텍스트마이닝 분석을 실시하였다. 이에 수집한 내용에 도출한 키워드 상위 50개의 핵심단어를 UCINET 6 네트워크 분석프로그램을 통해 단어의 연결중심성과 전체 네트워크들의 중심성을 계산하여 사회연결망 관점에 따라 태권도원에 관련된 단어들의 관계구조를 분석하고, 동시에 발생한 빈도의 관련성을 파악하였다. 또한, 태권도원과 관련된 단어들의 연결망을 보기 위해 UCINET 6 내에 Netdraw 기능을 이용하여 네트워크 분석을 시각화하였고, 군집을 도출하기 위하여 CONCOR분석을 하였다.

III. 데이터 분석 결과

1. 태권도원에 대한 키워드 분석결과

태권도원 핵심 키워드 상위 50개의 텍스트마이닝을 실시한 결과, 추출된 단어의 빈도와 연결중심성, 네트워크 중심성, 밀도는 <표 4>와 같다. 공출현 빈도가 높은 상위 10개의 키워드는 태권도(18,048건), 무주군(13,845건), 전라북도(5,969건), 세계(4,521건), 스포츠(3,921건), 공연(3,265건), 태권도진흥재단(2,610건), 한국(2,555건), 체험(2,523건), 사람(2,503건) 순으로 나타났다.

연결중심성이란 네트워크 내에 단어들이 태권도원이라는 단어의 중심에 얼마나 가까운지 나타내는 지표로서, 연결중심성 값이 크면 클수록 네트워크 내의 영향력이 높다(홍지숙, 오익근, 2016). 본 연구에서 키워드의 연결중심성

값이 큰 상위 10개 키워드는 태권도(0.064), 무주군(0.057), 전라북도(0.017), 공연(0.016), 체험(0.015), 아이(0.015), 진행(0.014), 세계(0.014), 사람(0.012), 모노레일(0.011) 순으로 나타났다. 추가적으로 네트워크 내의 전체 노드와 노드 사이의 관계 범위를 나타내는 네트워크 중심성은 20.279%로 나타났다.

표 4. 태권도원에서 추출된 단어의 빈도와 연결중심성

순위	단어	빈도	연결중심성	순위	단어	빈도	연결중심성
1	태권도	18048	0.064	26	만들다	1497	0.011
2	무주군	13845	0.057	27	방문	1481	0.010
3	전라북도	5969	0.017	28	중국	1469	0.004
4	세계	4521	0.014	29	참가	1440	0.007
5	스포츠	3921	0.006	30	북한	1433	0.002
6	공연	3265	0.016	31	성지	1422	0.004
7	태권도 진흥재단	2610	0.011	32	선수	1404	0.008
8	한국	2555	0.010	33	아이	1369	0.015
9	체험	2523	0.015	34	국기원	1347	0.009
10	사람	2503	0.011	35	수련	1335	0.006
11	이벤트	2495	0.007	36	공간	1324	0.008
12	대회	2182	0.011	37	올림픽	1297	0.004
13	진행	2177	0.014	38	운영	1271	0.010
14	시범	2049	0.009	39	전주	1224	0.003
15	프로그램	1902	0.010	40	태권도인	1180	0.004
16	동영상	1897	0.005	41	박물관	1172	0.009
17	역사	1879	0.005	42	설천면	1156	0.003
18	가족	1874	0.008	43	경기	1121	0.007
19	교육	1830	0.007	44	경기장	1095	0.007
20	열리다	1745	0.010	45	청소년	1070	0.004
21	개최	1717	0.009	46	사용	1069	0.005
22	전망대	1674	0.011	47	즐기다	1064	0.008
23	여행	1656	0.008	48	해외	1023	0.003
24	행사	1656	0.010	49	축제	1023	0.005
25	모노레일	1500	0.011	50	관람	1021	0.007

네트워크 중심성(Network Centralization) = 20.279%

2. 태권도원에 대한 N-gram 네트워크 분석결과

네트워크 내에서 연결된 관계를 시각적으로 도출하기 위하여 N-gram 네트워크 그래프를 분석하였다. 분석 결과 상위 10개 관계도는 전라북도-무주군(2,189), 무주군-설천면(985), 시범-공연(926), 태권도-시범(767), 태권도-성지(749), 세계-태권도(724), 세계-태권도인(643), 라이프-스포츠(585), 북한-세계(492), 스포츠-사람(460) 순으로 나타났다. 결과는 <표 5>와 <그림 1>과 같다.

표 5. 태권도원 N-gram 네트워크 분석결과

순위	키워드	키워드	빈도
1	전라북도	무주군	2189
2	무주군	설천면	985
3	시범	공연	926
4	태권도	시범	767
5	태권도	성지	749
6	세계	태권도	724
7	세계	태권도인	643
8	라이프	스포츠	585
9	북한	세계	492
10	스포츠	사람	474



그림 1. 태권도원 N-gram 네트워크 분석결과

3. 태권도원에 대한 네트워크 시각화 분석결과

네트워크 내에 나타난 패턴과 연결정도의 관계를 도출하기 위하여 연결중심성을 바탕으로 네트워크를 시각화한 분석 결과는 <그림 2>와 같다.

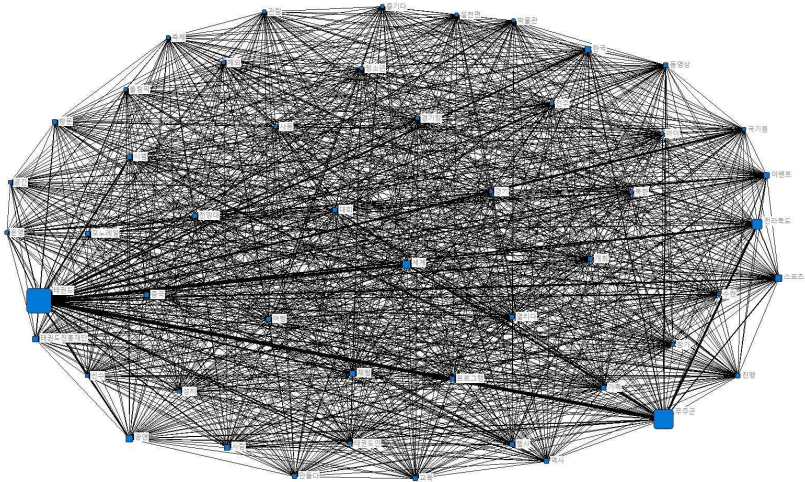


그림 2. 태권도원 네트워크 시각화 분석결과

4. 태권도원에 대한 CONCOR 분석결과

태권도원에 대한 CONCOR분석을 실시한 결과 ‘수련 및 교육장소’, ‘방문체험 콘텐츠’, ‘지자체 속성’, ‘가치속성’, ‘대회속성’으로 총 5가지 영역의 그룹으로 형성되었으며 이는 <그림 3>과 같다.

첫째, 공간, 시범, 운영, 경기, 수련, 국기원과 같이 태권도원에서 진행되는 수련 프로그램과 장소에 관한 단어들이 추출됨에 따라 ‘수련 및 교육장소’로 명명하였다.

둘째, 태권도, 태권도인, 성지, 올림픽, 선수, 한국, 방문, 가족, 태권도진흥재단, 역사, 사람, 세계와 같이 태권도의 성지로서 태권도인의 인식과 태권도원이 가진 속성들을 표현할 수 있는 내용이 추출됨에 따라 ‘가치속성’으로 명명하였다.

셋째, 무주군, 축제, 행사, 개최, 교육, 해외, 청소년과 같이 지역 인프라를 표현하는 단어들이 추출됨에 따라 ‘지자체 속성’으로 명명하였다.

넷째, 박물관, 관람, 경기장, 즐기다, 체험, 전망대, 프로그램, 아이, 공연, 모노레일과 같이 태권도원을 방문하여 경험할 수 있는 체험 콘텐츠들과 장소들이 추출됨에 따라 ‘방문체험 콘텐츠’로 명명하였다.

다섯째, 스포츠, 열리다, 참가, 북한, 대회, 중국과 같이 태권도 경기 대회에 관한 단어와 국가들이 추출됨에 따라 ‘대회속성’으로 명명하였다.

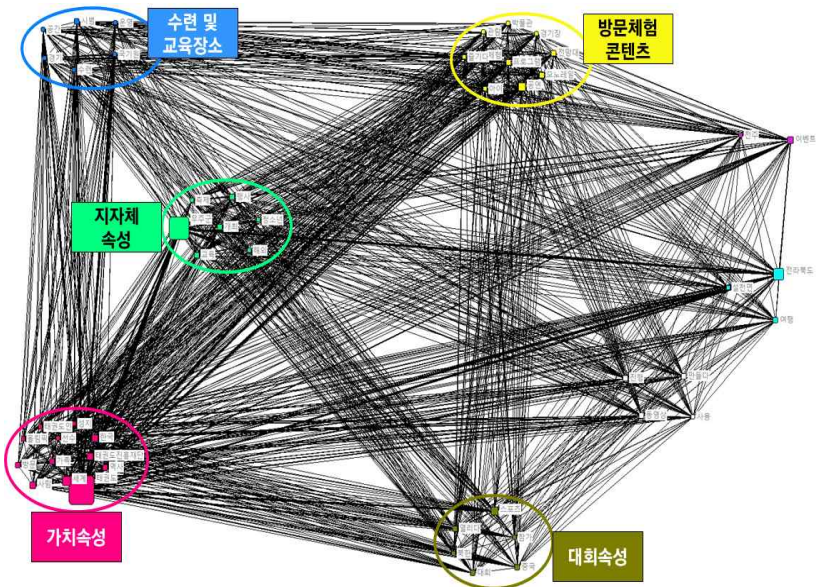


그림 3. 태권도원 CONCOR 분석

IV. 논의

본 연구는 빅데이터 분석 기법 중 텍스트마이닝 기법을 사용하여 세계 태권도인의 성지인 태권도원의 키워드들을 살펴보았다. 이에 본 연구의 CONCOR 분석 결과를 토대로 한 논의는 다음과 같다.

첫째, 태권도원의 상징적 가치와 공간을 나타내는 ‘수련 및 교육장소’로 분류되었다. 이와 관련된 키워드 빈도와 연결중심성은 20위에서 30위의 중위권에 있으며, N-gram의 경우 시범과 공연이 3위로 나타났다. 이렇듯 키워드로 나타난 원인으로는 태권도원을 방문하는 근본적인 이유로 판단된다. 태권도원은 본래 태권도인의 수련 장소 및 성지로 만들어졌으며, 국기원은 노후화되었고 규모가 작은 만큼 넓은 부지에서 전 세계 태권도인이 수련할 수 있는 공간을 제공하기 위해 만들었다. 정경희, 김범용, 유광길(2017)은 태권도원의 경쟁력요인을 갖추는 방안에 관한 연구를 통해 소비자인 시설이용자나 충성도 강한 고객을 확보하지 못한다면 발전과 성장을 기대하기 어려우며, 태권도원의 정체성이 부족해진다고 하였다. 또한, 정경희, 황보선, 유광길(2015)은 태권도원의 효율적인 운영관리에 관해 연구하였는데, 전체 17개 항목 중 3위 시설의 공간활용과 4위 체계적인 교육 훈련이 중요한 요소로 작용한다고 하였다. 따라서 태권도인의 수련 및 교육의 장소는 태권도원에 매우 중요한 근본적인 키워드라고 할 수 있다. 과거 서울 강남에 있는 국기원에서 사범교육 및 심판교육을 받았지만, 현재는 전북 무주에 있는 태권도원에서 이루어지고 있다. 국기원의 수용인원 부족으로 태권도원으로 변경되었지만, 무주는 이동하기에 멀고, 시간이 오래 걸리기 때문에 태권도인마저도 교육받기를 기피하고 있다. 서울에서는 무주 태권도원까지 가는 교통편이 부족하고 시간대를 맞추기 매우 어렵다. 그뿐만 아니라 국기원에서 받는 교육과 태권도원에서 받은 교육의 질적인 차이를 느끼지 못하고 있다. 이에 매년 연수와 교육을 받을 때면 국기원에서 교육을 받기 선호하는 현상이 벌어지고 있다. 태권도원은 본래 태권도인이 수련하고 세계적인 대회를 개최하며 태권도의 위상

을 높이기 위해 만들어진 공간이다. 따라서 태권도인이 방문하고 싶은 수련 공간, 배움이 넘치는 공간을 조성하고, 실제 태권도 수련프로그램을 자세히 살펴어 세부적인 수정을 거쳐야 할 것으로 판단된다.

둘째, 태권도원의 근본적인 가치와 내포된 의미를 나타내는 ‘가치 속성’으로 분류되었다. 이와 관련된 10위 내 키워드 빈도는 1위 태권도, 4위 세계, 7위 태권도진흥재단, 10위 사람으로 도출되었으며, 연결중심성은 1위 태권도, 8위 세계, 9위 사람으로 나타났다. 또한, N-gram 결과 5위 태권도-성지, 6위 세계-태권도, 7위 세계-태권도인으로 나타났다. 이러한 원인은 태권도원이 상징하는 의미들과 앞으로의 지향점에 관한 키워드로 제시된 것을 알 수 있다. 김상태 외 7인(2014)의 좌담회에서 태권도원은 태권도 종주국의 성지로서 구체적인 조성을 통해 태권도 허브 및 상징성을 강화하고, 국내외 홍보마케팅 펼쳐 세계적인 태권도 명소로 발전되어야 한다고 하였다. 또한, 이정희, 최문용(2017)은 태권도원의 발전 방향을 제시하였는데 첫째, 태권도는 폐쇄성이 강하여 국제적으로 공감된 운영이 부족하고, 둘째, 내한방문객들이 볼 수 있는 태권도원 관련 안내자료조차 없으며, 공항 내에 비치되어 있지 않아 개선책이 필요하다고 하였다. 나아가 세계태권도연합뉴스(2015)는 중국이 태권도를 자신의 무술이라고 주장하기 때문에, 태권도의 역사성을 보존하기 위해서는 태권도원을 방문한 사람들이 이목을 이끌 수 있는 구조물을 설치해야 한다고 하였다. 따라서 태권도진흥재단은 태권도원을 세계적인 태권도의 성지로서 해외 방문객에게 널리 알리고, 태권도의 역사성과 정체성을 갖추어야 할 것으로 판단된다. 태권도원은 곧 태권도를 대표하는 장소로서 그 가치가 매우 높기 때문이다. 중국의 소림사는 495년에 창건되어 2010년 유네스코 세계유산으로 등재되었고, 그 역사가 매우 깊어 현재 중국 정부의 공인된 체육 교습 기관으로 사용되고 있다. 태권도원은 소림사에 비하면 역사가 매우 짧지만, 지금부터라도 후세를 위하여 태권도원을 보전하고, 정부가 태권도의 가치를 이해하고 투자를 한다면 소림사를 뛰어넘을 태권도의 성지로 거듭날 것으로 판단된다. 문화체육관광부(2019. 12. 6.)의 보도자료에서 태권도 가치 회복을 통해 세계경쟁력 강화하기 위하여 ‘태권도 대사범(Grand Master)’을

지정하여 수련 프로그램운영과 태권도를 통한 심신 단련을 하는 ‘태권스테이’를 시행한다고 하였다. 이러한 아이디어를 지속해서 발전시켜 진정한 태권도의 성지로 거듭나도록 하여야 한다.

셋째, 지역의 속성 및 홍보 효과를 나타내는 ‘지자체 속성’으로 분류되었다. 이와 관련된 키워드 빈도와 연결중심성의 순위는 2위 전라북도와 3위 설천면으로 나타났고, N-gram 결과 1위 전라북도-무주군, 2위 무주군-설천면으로 나타났다. 이러한 원인은 태권도원의 위치가 전라북도 무주군 설천면에 소재해 있고, 태권도원 위치의 검색 및 업로드 자료들을 통해 홍보효과가 나타난 것을 알 수 있다. CONCOR 분석을 통해 자세히 살펴보면, 무주군에서 준비했던 축제, 행사, 개최, 청소년, 해외와 같이 태권도원과 연관된 행사와 이벤트 관련 단어가 추출된 것을 알 수 있다. 한국문화정보원(2019)의 연구에서는 무주 태권도원을 선정하여 지자체 체험관 활성화 방안을 제시하였고, 이도우, 이재봉(2015)의 연구에서는 무주 태권도원을 중심으로 전라북도 관광산업의 부흥과 지역경제 활성화를 위해 전북 관광 중점도시화를 위한 특별구역 지정이 필요하고, 나아가 지속적 관광 개발을 위한 정부의 제도적인 장치가 필요하다고 하였다. 또한, 김범용, 정경희, 유광길(2019)은 태권도원의 지역 사회 발전을 위해서는 긍정적인 이미지 형성이 매우 중요하며, 이를 해결하기 위해 태권도원의 일률적인 방문객뿐만 아니라 지자체와 함께 지역을 홍보하고 태권도원을 통하여 지역을 알리는 노력이 필요하다고 하였다. 따라서 전라북도, 무주군의 지자체와 태권도원의 발전을 위해서는 서로 상호협력하는 방안이 필요할 것으로 판단된다. 우선 서울과 태권도원의 교통편의를 위해서 지자체에서는 수도권 서틀버스를 운영하여 접근성을 높이고, 교통의 편리성도 높여야 한다. 태권도원은 유소년의 방문객이 많기 때문에 일반 대중 교통 및 고속버스로는 거리와 시간, 금전적인 부분이 부담으로 작용한다. 또한, 지자체에서는 버스에 무주군의 홍보물을 부착하여 외부광고를 하고, 스키장 서틀버스터럼 시간상으로 그다지 멀지 않다는 것을 홍보한다면 다양한 효과를 볼 수 있을 것으로 판단된다. 또한, 태권도원에서 대회를 개최할 시에는 지자체의 특수성을 잘 나타내는 특산물 등을 판매하여 무주군의 경제적 이익

및 홍보 효과를 높일 수 있다.

넷째, 태권도원을 방문하여 체험하고 느낄 수 있는 장소들을 나타내는 ‘방문체험 콘텐츠’로 분류되었다. 이와 관련된 키워드 빈도와 연결중심성은 공연, 체험, 전망대, 모노레일, 아이, 박물관, 관람, 즐기다, 경기장으로 대부분 빈도 순위보다 연결중심성 순위가 높다. 즉, 단순히 빈도수가 높은 것이 아닌 실제 태권도원과 관계가 높은 키워드라고 할 수 있다. 이는 방문객들이 정말로 필요한 자료들을 검색할 때 나타나는 키워드이다. 이러한 키워드는 태권도원을 방문하는 이유로 설명할 수 있는데, 방문하였을 때 어떠한 것을 체험하고 제공 받을 수 있는지를 확인할 수 있다. 장용석(2019)의 연구에서 태권도원의 홈페이지는 단순히 시설소개 및 위치만 소개되어 있어 일방적인 정보전달만 되기 때문에 일반 관광객을 대상으로 태권도원을 체험시키기에는 무리가 있다고 하였다. 또한 한국문화정보원(2019)의 연구에서는 태권도원의 대표 콘텐츠 구축 및 체험관 활성화 방안 모색이 필요하며, 이를 해결하기 위해 13세 이하의 유아나 초등생을 타겟층으로 정하고, 부모님과 함께할 수 있는 체험 콘텐츠 및 부모님 편의시설 구축이 필요하다고 하였다. 또한, 문화체육관광부(2019. 12. 6.)의 보도자료에서는 태권도원 체험연수 프로그램을 20인 이상의 단체예약만 받았지만, 개인, 가족 등의 소규모 단위도 예약할 수 있도록 확대 운영한다고 하였으며, 나아가 향후에는 태권도를 종합적으로 체험하는 복합체험시설을 구축하여 한국문화 체험캠프 운영 등 다양한 프로그램을 운영한다고 하였다. 따라서, 이미 구성된 체험 시설 및 프로그램을 유지하되 고객층을 세분화하고, 품새, 시범, 겨루기 모두를 접할 수 있는 다양한 체험관을 건설하여 방문 만족도를 높여야 할 것으로 판단된다.

다섯째, 태권도원에서 개최한 대회의 관련 키워드를 나타내는 ‘대회속성’으로 분류되었다. 이와 관련된 키워드 빈도와 연결중심성은 5위 스포츠, 12위 대회, 20위 열리다, 29위 참가 등이 상위 키워드로 나타났으며 N-gram 결과 10위 스포츠-사람이 도출되었다. 이는 태권도원에서 세계태권도대회를 개최하여 나타난 그룹이라 판단된다. 김효종(2017)과 최창운(2017)은 ‘2017 무주 세계태권도선수권대회’로 얻은 이미지를 바탕으로 무주군 민자사업을

추진하고, 태권도원이 진정한 세계태권도의 성지로 거듭날 수 있다고 하였다. 또한, 세계 대회를 통해 태권도원의 가치가 올라가고, 강력한 시장성을 보여주었다(박종화, 2018). 따라서, 태권도원은 세계적인 태권도대회를 유치하기 위한 노력들이 필요할 것으로 판단된다. 테니스, 골프, 레이싱 등 각 종목의 세계적인 대회들에는 지역 명칭이 노출되기 마련이다. 즉 태권도에서도 세계적인 4대 대회를 형성하여 무주 태권도원이 그중 대표하는 대회로 거듭나야 한다. 현재 겨루기 부분에서 월드 태권도 그랑프리 대회가 있는데 이러한 대회들을 서둘러 선점하여 대회를 활용한 지역 마케팅에 노력을 기울여야 한다. 그뿐만 아니라 품새와 시범과 같이 아직 세계대회가 미비한 종목들에 대해 무주에서 개최하도록 노력해야 한다. 이에 세계태권도연맹(WT)과의 밀접한 유대관계를 통해 태권도원의 가치를 표출해야 한다.

V. 결론 및 제언

1. 결론

본 연구는 태권도원에 대한 텍스트마이닝 기법을 통해 ‘상위 50개 키워드’, ‘연결중심성’, ‘N-gram 네트워크’, ‘CONCOR 분석’과 같은 결과를 도출하였다.

첫째, 태권도원의 상위 50개 키워드 분석결과 상위 10개 키워드로는 태권도, 무주군, 전라북도, 세계, 스포츠, 공연, 태권도진흥재단, 한국, 체험, 사람 순으로 나타났다.

둘째, 태권도원의 연결중심성이 높은 상위 10개 키워드로는 태권도, 무주군, 전라북도, 공연, 체험, 아이, 진행, 세계, 사람, 모노레일 순으로 나타났다.

셋째, 태권도원 관련 키워드를 토대로 N-gram 네트워크 그래프 결과 상위 10개 키워드 관계로는 전라북도-무주군, 무주군-설천면, 시범-공연, 태권도-

시범, 태권도-성지, 세계-태권도, 세계-태권도인, 라이프-스포츠, 북한-세계, 스포츠-사람 순으로 나타났다.

넷째, 태권도원의 CONCOR분석 결과 ‘수련 및 교육장소’, ‘가치속성’, ‘지자체 속성’, ‘방문체험 콘텐츠’, ‘대회속성’으로 총 5가지 범주로 형성되었다.

이러한 결과를 토대로 태권도원이 개장한 이후에 나타난 6년간의 관련 키워드들을 토대로 방문객과 실제 태권도인이 느끼고 경험한 태권도원의 인식을 분석하였다. 이에 태권도원의 관계자들은 실제 태권도원을 방문하는 사람들이 좋아하고 관심을 두는 키워드들을 따로 분류하고, 좀 더 발전되고 개발해야 할 장소와 공간, 그리고 이벤트에 대한 노력을 기울여야 한다. 이에 고객이 실제 체험 후기나 방문 후기에 관한 사진 업로드를 할 수 있는 앱 개발을 지향하고, SNS를 활용한 차별화된 마케팅 전략들이 나온다면 좀 더 많은 사람이 선호하고 방문하고 싶은 공간으로 변화될 것이다.

2. 제언

결론을 종합하여 태권도원에 대한 다양한 키워드와 키워드 간의 관계, 그리고 범주영역 그룹까지 확인하였다. 이에 본 연구과정에서 나타난 문제점과 향후 연구에 대한 과제를 제시하고자 한다.

첫째, 과거의 데이터를 수집하는 과정에서 오래된 데이터들의 경우 자체 삭제되거나 변형될 수도 있어서 모든 내용을 정확히 얻고 파악하는 데 어려움이 있었다.

둘째, 본 연구에서 제시된 키워드를 토대로 향후 연구에서는 실제 방문객들의 경험하였던 감정과 느낌에 대한 단어들의 텍스트마이닝 연구를 조사한다면 좀 더 의미있는 복합연구가 될 것이다.

셋째, 본 연구는 설문조사를 진행한 것이 아닌 인터넷상에 있는 자료를 중심으로 조사하였기에 객관화된 자료라고 하기에는 다소 무리가 있다. 이에

본 연구를 토대로 설문 또는 면담을 통한 추가적인 자료수집이 된다면 좀더 심층적인 연구가 될 것이다.

넷째, 향후 연구에서는 SNS의 영역을 넓혀 다양한 종류의 매체(BAND, 인스타그램 등)를 추가 확보하여 연구를 진행한다면 좀 더 세부적인 내용을 파악할 수 있을 것이다.

참고문헌

- 김상태, 박노신, 박병직, 이봉, 이정우, 전효재, 양철모(2014). 좌담회-태권도원 활성화의 조건. **한국관광정책**, 58, 43-73.
- 김선수(2014). 태권도의 관광상품화를 위한 마케팅 전략 방안 : 한국 전통 문화를 중심으로. **무예연구**, 8, 1-15.
- 김종석, 정경희, 유광길(2014). 태권도원 내부고객 지원활동 시장지향성 및 고객지향성의 관계. **한국체육과학회지**, 23(5), 771-781.
- 김진기, 정경희, 유광길(2017). 태권도원 민간투자 활성화요인에 관한 우선순위 분석. **한국체육과학회지**, 26(1), 561-571.
- 김환표(2013). **트렌드 지식 사전**. 인물과사상사.
- 김효종(2017. 11. 17). **무주WTF 세계태권도선수권대회 결과보고회 “10000억 이상 경제 파급 효과”**. 전북일보. 2017. 11. 17. 검색, www.jjan.kr/news/articleView.html?idxno=1142641.
- 문화체육관광부(2019. 12. 6.). **일상에서 누리는 문화중심의 태권도**. 보도자료, 2020. 03. 19. 검색. www.mcst.go.kr/attachFiles/viewer/skin/doc.html?fn=20191206083446336009948909_TCS_SCHMNG20191206083519235745.hwp&rs=attachFiles/viewer/result/press.
- 박인준, 조성균(2018). 태권도 문화콘텐츠로서 융합 프로그램의 패러다임 인식과 실천의 내러티브 연구. **무예연구**, 12, 43-63.
- 박종화(2018). 빅데이터를 활용한 ‘2017년 무주 세계태권도선수권대회’ 탐색적 연

- 구: 소셜 미디어 자료를 중심으로. **국기원태권도연구**, 9(3), 77-96.
- 배진남(2013). **개원 준비 한창인 태권도성지...미리 가본 태권도원**. 연합뉴스, 2020. 03. 02. 검색. www.yna.co.kr/view/AKR20131122057600007.
- 세계태권도연합뉴스(2015. 1. 31.). **태권도원이 나아가야 할 방향**. 2020. 02. 28. 검색. www.wtu.kr/292.
- 양명모(2019). **디자인 사고 방법을 활용한 태권도원 문화콘텐츠 구축 방안**. 미간행 박사학위논문. 경희대학교 체육대학원.
- 양명모, 조성균(2019). 디자인 사고 방법을 활용한 태권도원 문화콘텐츠 구축 방안. **무예연구**, 13, 97-115.
- 유광길(2019). 태권도원 매력도 분석을 통한 효율적 전략 탐색. **한국사회체육학회지**, 75, 27-38.
- 윤동길(2017). **태권도원 활성화, 국기원 완전 이전·정부지원 절실**. 전민일보, 2020. 03. 10. 검색. www.jeonmin.co.kr/news/articleView.html?idxno=174074.
- 이도우, 이재봉(2015). 태권도원을 통한 무주지역 활성화 방안. **한국스포츠학회지**, 13(2), 269-281.
- 이미연, 전익기(2014). 결정적사건기법(CIT)을 통한 태권도원 시설에 대한 방문 전 기대, 방문 후 성과 및 개선방안 연구. **한국체육학회지**, 53(5), 309-325.
- 이정학, 김옥기(2012). 인문·사회과학편: 태권도공원과 연계한 태권도 관광 상품 개발. **한국체육학회지**, 51(6), 179-188.
- 이정학, 이재문, 김옥기, 김형근(2017). 빅데이터 텍스트마이닝 분석을 통한 수영복 인식에 관한 연구. **체육과학연구**, 28(1), 104-116.
- 이정희, 최문용(2018). 태권도원의 브랜드홍보 전략수립을 위한 기초연구. **브랜드디자인학연구**, 16(2), 237-246.
- 장용석(2019). 관광스토리텔링을 활용한 무주 태권도원 관광콘텐츠 개발. **국기원태권도연구**, 10(1), 107-126.
- 장혜영, 박현진, 최승희(2016). 무주 태권도원 문화관광상품 개발 연구. **디지털융복합연구**, 14(2), 351-357.
- 전채남, 서일원(2013). 빅데이터 분석의 기술마케팅 활용에 관한 연구: 잠재 수요기업 발굴을 중심으로. **마케팅논집**, 21(2), 181-203.

- 정경희, 김범용, 유광길(2017). 태권도원 경쟁력요인에 관한 계층구조 분석. **한국체육과학회지**, 26(3), 499-510.
- 정경희, 황보선, 유광길. (2015). 태권도원의 효율적인 운영관리를 위한 AHP 분석. **한국체육과학회지**, 24(5), 907-917.
- 차민경(2015). 국내 언론에 나타난 ‘예술경영’ 관련 이슈의 의미연결망 분석: 1990년부터 2014년까지 국내 일간지 기사 분석을 중심으로. **문화정책논총**, 29(2), 168-200.
- 차민경, 권상희(2015). 언론의 ‘창조경제’에 대한 의제설정 의미연결망 분석. **한국언론학보**, 59(2), 88-120.
- 최창윤(2017. 11. 18.). 무주군, 2017 무주 WTF 세계태권도선수권대회 결과보고회 가제. 뉴스메이커. 2017. 11. 18. 검색. www.newsmaker.or.kr/news/articleView.html?idxno=52515.
- 최천. (2013). 태권도의 사회적 공유가치. **무예연구**, 7, 1-13.
- 태권도진흥재단(2020). 태권도원 방문객 현황. 2020. 03. 09. 검색. www.tpf.or.kr/tpf/contents/tpf0406.do?&schM=list&act=UPDATE&page=1&viewCount=10&schBdcode=&schGroupCode=.
- 한관중(2003). 사회과학 방법론으로서의 연결망 분석기법 적용의 의의와 연구과제: 의미와 연결망 분석 (semantic network analysis) 을 중심으로. **사회과교육연구**, 10, 219-235.
- 한국문화정보원(2019). 지자체 체험관 활성화 방안 -무주 태권도원 체험관 중심으로-. **한국지방정부학회지**, 11, 637-657.
- 홍준기(2016). 소셜미디어를 통한 태권도원 홍보 전략에 관한 연구. 미간행 석사학위논문. 한국체육대학교 사회체육대학원.
- 홍지숙, 오익근(2016). 소셜미디어 빅데이터를 활용한 항공사 이미지 변화 분석: K항공의 램프리턴을 중심으로. **관광연구저널**, 30(6), 119-133.
- Brown, B., Chui, M., & Manyika, J.(2011). Are you ready for the era of 'Big data'. *McKinsey Quarterly*, 4(1), 24-35.

ABSTRACT

A Study On Analyze of ‘Taekwondowon’ Using Big data Text-mining Analysis

Jeong, Kwon-Hyuk, Jeon, Ik-Ki(Kyunghee Univ.)

The purpose of this study is to identify the keywords related to Taekwondowon, the holy land of Taekwondo people, and to provide basic data on the development direction of Taekwondowon. This research is based on data by using Daum and Naver web sites, blogs, news, Cafes, JISIKiN / TIP and Web and News offered by Google. The data collection period was limited to 6years from April 24, 2014 to April 16, 2019. The results of anlysis are ad follows: First, Collected Data are 8,283 cases. Second, Key words related to Taekwondowon are those - Taekwondo, Muju-gun, Jeollabuk-do, World, Sports etc. Third, the median of connectivity values are showed high in Taekwondo, Muji-gun, Jeollabuk-do, Show, Experience etc. The total network connectivity 20.279%. Forth, as a result of conducting N-GRAM network graph are those - Jeollabuk-do-Muju-gun, Muju-gun-Seolcheonmyeon, Demonstration-Experience, Taekwondo-Demonatration, Taekwondo-Holy land etc. Fifth, As a result of CONCOR analysis, 5 groups are formed - ‘Training and Education Place’, ‘Visiting experience content’, ‘Local property’, ‘Value property’, and ‘Competition property’.

Key words: Taekwondowon, Big data, Text-mining

논문투고일 : 2020.03.23.

심 사 일 : 2020.04.10.

심사완료일 : 2020.04.30.