

# 소셜미디어에 공유한 정보를 통한 개인 성격유형 분석 앱 개발

## Development of Personal Character Analyzing Application Based on the Opened Information at the Social Media

한정화\*, 박진완\*\*  
트리플\*, 중앙대학교 첨단영상대학원\*\*

Jung hwa Han(jurohan@empal.com)\*, Jin wan Park(jinpark@cau.ac.kr)\*\*

### 요약

스마트폰의 빠른 보급과 소셜 미디어의 급격한 성장에 힘입어 온라인 상에 기록되고 있는 개인의 정보의 양 또한 급격히 증가하고 있다. 이러한 환경의 변화에 따라 사용자들이 생성한 정보를 분석하여 새로운 정보를 얻기 위한 노력이 증가하고 있다.

기존의 성격유형 분석 방법의 경우 자기 보고 문항을 통해 사용자의 유형을 탐색하는데, 이는 검사 대상자의 정보들을 심리학 전문가가 확보하기 어려웠기 때문이다. 하지만 유명인의 경우 그들의 정보를 쉽게 취득할 수 있고 이러한 정보를 통해 해당 인물의 성격 유형을 분류하는 연구들이 다양하게 이뤄진다는 점에서 볼 때, SNS 에 기록된 개인 사용자들의 정보를 통해 성격유형 분석이 가능할 것이라는 가설을 세울 수 있었다. 본 연구에서는 이를 증명하고자 분류지표를 설계하여 페이스북 네이티브 앱을 개발하였고, 이를 통해 사용자가 소셜 미디어 상 공유한 정보에는 정보 소비 기호가 반영되어 있으며, 이를 통해 성격유형을 파악할 수 있음을 확인할 수 있었다.

■ 중심어 : | 소셜 데이터 활용 | MBTI 지표 활용 | 페이스북 앱 |

### Abstract

The amount of personal information has increased dramatically due to the prevalent use of smartphones and the rapid growth of social networking services. Under these circumstances, there has been a lot of efforts to obtain new information based on the overflowing personal data.

The conventional character analysis method, which heavily relies on personal surveys, had some limitations in that it was difficult for psychologist to have a complete access to the surveyed results. When it comes to celebrities, however, it is relatively easy to access to their information through various media. Therefore there has been various researches that examined celebrities' personalities. On the contrary, not many studies have focused on analyzing the characteristics of the general public whose information is not so accessible. In this research, we suggest a method to analyze ordinary people's characteristics based on information available via their social networking services. This research focuses on developing a Facebook-native application, which examines the user's character type based on the posts shared in the user's Facebook page.

■ keyword : | Use of Socail Data | Use of MBTI Data | Facebook App |

## I. 서론

### 1. 연구 배경 및 목적

최근 한국의 스마트 폰 보급률은 빠르게 증가하여 현재 73%를 넘으며[1] 개인이 온라인 채널과 연결되는 시간 또한 빠른 속도로 증가하고 있으며, 이러한 환경에 기반해 대표적인 소셜 미디어인 페이스북 북의 국내 사용자수가 1,000만명이 넘어서고 있다[2]. 또한 카카오톡 등 모바일에 특화되어 개발된 서비스와 그 연계 서비스들 또한 빠른 속도로 사용자를 확보하고 있다. 대표적인 예로 선데이 토츠의 애니팡은 카카오톡의 연계 서비스로 2012년 7월 처음 공개된 후 1년이 채 안되어 3,000만 명이 넘는 이용자를 만들기도 하였다[3]. 이러한 변화는 기존의 매스 미디어인 라디오, TV 등이 전세계 사용자 5억 명을 넘어서는 데 걸린 시간이 10년 이상의 오랜 시간이 걸린 것과는 비교할 수 없는 급격한 변화임이 분명하다.

이러한 환경에서 소셜 미디어 이용자 간에 소통의 과정에서 만들어지는 정보들은 그 정보의 범위와 양적인 측면에서도 빠른 속도로 확대되고 있다. 또한 이와 같은 환경의 변화에서 만들어지는 개개인의 자료들은 분석이 용이한 디지털 데이터의 형태를 띠고 있으며, 오늘날 이러한 데이터를 소유한 대부분의 기업들은 자신들에게 축적된 데이터를 기업 외부에서 Open API를 통해 새로운 결과물을 만들 수 있는 환경을 제공하고 있다. 이는 개개인이 만들어 내는 데이터의 분석을 통해 고객들이 기업 또는 제품에 대해 갖는 생각을 확인하여 구체적인 요구사항을 파악할 수 있다는 가능성이 존재하기 때문에 현재 그 중요성이 점차적으로 커지고 있다.

본 연구에서는 사용자들이 페이스북 북에 공유한 온라인상의 정보들을 MBTI 성격유형 분석 지표를 활용하여 개발된 분석 DataBase에 대입하여 이용자의 성격유형을 도출하는 모델을 적용한 Facebook App을 개발하였다. 본 연구에서는 URL이라는 단일 지표를 통해 분석을 진행하기 때문에 감각(Sensing)과 직관(iNtuition), 사고(Thinking)와 감정(Feeling) 이라는 4개의 분류 기준에 사용자의 공유 정보를 대입하여 1차 분석을 진행하고, 도출된 1차 분석 결과에 기준해 이용

자가 공유했던 URL을 한번 더 분석하여 최종 결과가 도출되는 형태로 분류 모델을 구성하였다. 이후 본 모델의 이용자들이 자유롭게 결과에 대한 의견을 달 수 있는 영역을 구성하였다. 이를 통해 이용자들이 본 모델을 통해 도출된 결과에 대한 긍정적 혹은 부정적인 반응에 대해 확인할 수 있었다.

## 2. 이론적 배경

### 2.1 소셜 미디어

소셜 미디어는 소셜 네트워크의 기반 위에서 개인의 생각이나 의견, 경험, 정보 등을 서로 공유하고 타인과의 관계를 생성 또는 확장 시킬 수 있는 개방된 온라인 플랫폼을 의미한다. 기본적으로 양방향성을 활용하여 이용자들이 자발적으로 참여하고 정보를 공유하며 콘텐츠를 만들어 나가는 특성이 있다는 점에서 기존 전통적인 단방향의 미디어(TV, 신문, 잡지, 라디오 등)와는 구별되며, 이용자들이 콘텐츠 소비자임과 동시에 콘텐츠 생산자가 되는 것을 가능케 한다는 특징을 갖는다[4].

소셜 미디어는 온라인 커뮤니티와 유사한 측면도 있지만, 온라인 커뮤니티가 공통의 관심사를 가진 사람들이 특정 사이트에 모여 그룹을 형성한다는 특징을 가진 반면, 소셜 미디어는 개인을 중심으로 개인간 연결된 관계가 누적됨으로써 더 큰 네트워크를 형성할 수 있다는 점에서 차이가 있다[5].

또한 사용자의 사회적 활동 혹은 관계에 기반을 두었기 때문에 실제 사용자의 보편적 추구 가치, 성향에 기반해 정보 소비와 생성 활동이 이뤄진다는 특징이 있다.

### 2.2 소셜 미디어 분석

인터넷 이용자들은 소셜 미디어, 이메일, 메시지, 설문조사 등 많은 방법을 통해서 방대한 양의 비 구조적 텍스트 데이터들을 생성하고 있으며 최근 그 양이 급속도로 늘어나고 있는 상황이기 때문에 소셜 미디어 분석 분야가 많은 주목을 받고 있다.

대표적 연구 방향은 소셜 미디어를 통해 생성되고 유통되는 기록물을 통해 정보를 획득하고 이해하는 능력과 더불어 변화하는 사회에서의 적응 및 대처하는 능력을 의미하는 디지털 리터러시를 문화적 관점에서 연구

하는 방향이 있으며, 소셜 미디어 상의 데이터를 기존 데이터베이스 관리 도구의 데이터수집, 저장, 관리, 분석하는 역량을 넘어서는 비정형 데이터 셋으로 규정하고, 방대한 비정형 데이터에서 유용한 정보나 패턴을 추출하는 방향이 존재한다.

기존 연구 사례로는 사회연결망의 연결정도를 활용하여 상품 추천의 정확성을 향상시키는 방안에 대한 연구[6], 관계 네트워크의 정보 확산에서 정보 공유에 참여하는 혁신자와 허브가 정보의 이른 도약과 도약 시점의 확산시기에 미치는 영향에 대한 연구[7], 하이퍼 링크 데이터를 이용한 과학 커뮤니케이션 분석에 대한 내용을 다룬 사회 네트워크 이론과 방법론에 대한 연구[8] 등이 있다.

소셜 미디어 등에서 만들어지는 데이터의 분석은 특정 정보 혹은 사용자, 집단을 기준으로 하여 해당 정보의 공유에 참여한 사용자간의 관계 혹은 특성을 사회과학적·통계학적 분석방법을 통해 파악하고, 이를 통해 과거에는 불가능했던 사회구성원들에게 맞춤형 정보를 제공하거나 현대사회를 보다 정확하게 예측하고 효율적으로 작동하도록 정보를 제공할 수 있다는 점에서 가능성이 점차되고 있다[9].

### 1.3 분류(Classification) 기법

소셜 미디어와 같은 대규모의 정보 안에서 체계적이고 자동적으로 통계적 규칙이나 패턴을 찾아내는 기술을 데이터 마이닝(Data mining)이라고 하며, 이에 대한 세부 분야로는 분류(Classification), 군집화(Clustering), 연관성(Association), 연속성(Sequencing), 예측(Forecasting) 등이 있다[10].

이중 분류 기법은 이미 알려진 상호 배반적인 몇 개 그룹에 대해 각각의 그룹이 어떤 특징을 가지고 있는지 분류 모델을 만들고, 소속 그룹이 알려지지 않은 새로운 관측치가 어떤 그룹에 분류 될 것인가를 결정하는 다변량 분석방법이다. 분류 분석의 주요 목적은 새로운 관측치를 특정 그룹으로 할당하기 위한 최적의 분류규칙 (또는 함수, 모형)을 유도하는 것이다[11].

본 연구는 사용자가 공유한 정보 URL을 통해 MBTI의 16가지 성격유형 그룹에 분류한다는 점에서 분류 기

법을 응용하여 구성하였다.

### 1.4 MBTI 성격유형 분석 지표의 다양성 활용

MBTI(Myers-Briggs Type Indicator)는 인식과 판단에 대한 스위스심리치료가 융(Carl Jung)의 심리적 기능 이론, 그리고 인식과 판단의 방향을 결정짓는 융의 태도 이론을 바탕으로 제작되었다[12].

MBTI유형에 대한 분석은 자기보고(Self Report) 문항을 통해 대상자의 성격유형을 파악할 수도 있지만, 관련 연구를 수행한 심리학 전공자들의 경우 16가지 결과 유형별 특징에 기반하여 인물 혹은 캐릭터의 역할과 행동 패턴 등을 기준으로 대상자의 선호 유형을 유추할 수 있기 때문에, 이와 관련된 다양한 연구가 진행되고 있다.

그 예로 3D 픽사 애니메이션에 등장하는 캐릭터에 대한 성격유형을 분석한 오시룡, 석혜정의 연구[13], 자기 탐색과 자기이해라는 공통점에 기반하여 만다라 그림의 표현 양식과 MBTI 선호지표의 유의미한 상관 관계를 분석한 연구[14] 등이 진행된 바 있다.

본 연구에서는 자기 보고 문항을 대신하여, 소셜 미디어를 통해 사용자가 공유한 정보들을 통해 사용자의 인식 및 판단에 대한 선호 경향을 찾는 것에 목적이 있다. 또한 이 과정에서 수집된 정보 분석을 통해 현재 소셜 미디어를 통해 공유되는 정보 유형의 특징을 추가적으로 발견할 수 있었다.

## II. 연구 방법 설계 및 실험 진행

본 연구에서의 내용은 다음과 같은 분석 흐름도의 절차에 따라 수행 하였다. [그림 1]의 각 과정별 내용은 2.1 연구 진행 절차에서 자세하게 설명한다.

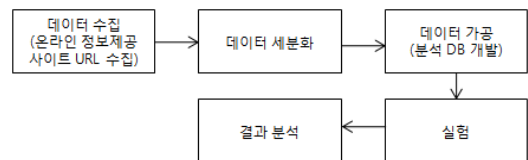


그림 1. 연구 흐름도

## 1. 연구 진행

### 1.1 데이터 수집

본 연구에서는 사용자가 공유한 정보URL별 분류 지표를 구성하지 않고, 상위 도메인, 즉 사이트 혹은 카테고리 단위에서 분류하는 방식으로 설계하였다. 그 이유는 현재 소셜 미디어 상에 공유되는 개별 정보들 중 복수 공유자가 존재하는 정보의 추출에는 한계가 있으며, 사용자들은 자신이 소비하는 정보의 성격, 기호에 따라 사이트에 접근하기 때문에 상위 개념인 사이트 혹은 정보 카테고리 단위에서도 해당 사이트의 정보를 공유하는 사용자 집단의 유형은 구분할 수 있기 때문이다. 예를 들어, IT와 관련된 정보를 제공하는 사이트에 접근하여 정보를 공유하는 사용자는 진실과 사실에 관심을 갖고 논리적이고 분석적이며 객관적으로 판단하는 특징을 갖는 사고형(Thinking) 유형일 가능성이 높으며, Design 과 Art에 관련된 정보를 공유하는 사용자는 사실 또는 사건 이면의 의미나 관계, 가능성을 더 잘 인식하는 직관(iNtuition) 유형일 가능성이 높다.

이러한 가설에 기반하여 연구의 첫 단계인 데이터 수집 단계에서는 국내외 정보제공 사이트에 대한 목록을 수집하였다. 국내의 경우 포털 사이트를 통한 정보 접근 및 소비가 일반화 되어 있기 때문에, 우선적으로 포털 사이트의 하위 뉴스/콘텐츠 제공 주소들에 대한 조사를 진행하였고, 이후 페이스북, 트위터 상 채널을 만들어 운영하는 정보 제공 사이트의 목록을 확인하였다. 그 결과 4개의 주요 포털사이트, 7개의 신문사, 12개의 경영 정보 제공 사이트, 9개의 IT 정보 제공 사이트, 5개의 디자인, 4개의 미디어 사이트, 정보제공형 SNS 사이트 5개에 대한 정보를 수집할 수 있었다.

### 1.2 데이터 세분화

수집된 정보 제공 사이트의 규모에 따라 세부 카테고리가 다양한 경우, 하위 서비스 단위 도메인을 추가로 추출하는 데이터 정제과정을 거쳤다. 이를 통해 수집된 정보 URL을 더욱 세분화하였다. 이를 통해 77개의 정보 제공 사이트에 대한 정보를 정리할 수 있었다.

표 1. 정보제공 사이트 URL 세분화 예시

사이트	카테고리	URL
다음		www.daum.net
	미디어 다음	http://media.daum.net
	스포츠 다음	http://sports.media.daum.net
	무비 다음	http://movie.daum.net/movieDetail
	연예 다음	http://media.daum.net/entertain
	텔존	http://bbs3.telzone.daum.net
	TV팟	http://tvpot.daum.net/story
스타일	http://bbs.style.miznet.daum.net	

### 1.3 데이터 분류 모델 개발

#### 1.3.1 정보 분류 모델 설계

본 연구에서는 URL이라는 단일 분석 지표를 통해 분류를 진행하기 때문에, 결과의 정확성을 높이기 위해 2회에 걸친 분류를 진행한다. 1차 분류에서는 사용자의 개별 정보들의 1차 선호 성향((Sensing / iNtuition / Thinking / Feeling)을 확인한 후, 집계하여 1차 성향을 파악한다. 1차 선호 성향을 S/N/T/F로 설정한 이유는 MBTI의 4가지의 양극적 선호 경향 중 정보의 공유에 있어 직접적으로 관여하는 경향은 '정보 수집 및 인식 기능의 척도'인 감각 (Sensing) / 직관 (iNtuition)과 '판단과 결정의 척도'인 사고(Thinking) / 감정 (Feeling) 기능이 활용되기 때문이다.

위의 과정을 통해 확인한 사용자의 1차 선호 성향에 대한 파악이 된 후엔, 다시 사용자가 공유했던 개별 정보들에 2차 분류 DataBase에 대입하되, 분류 기준을 1차 분석에서 도출된 1차 성향에 기준하여 분류한 후 가장 많이 집계된 성향을 사용자의 최종 성향으로 선정하는 형태로 정보 분석 모델을 설계하였다.

#### 1.3.2 분류 모델 Database 구축

데이터 정제 과정을 통해 수집된 도메인의 정보를 MBTI 유형분석에 경험이 있는 4명의 상담심리사들에게 제공하고, 1차 선호성향 (S/N/T/F)을 분류 기준으로 정하여 각각의 도메인에서 제공하는 정보를 확인하게 한 후 가장 공유할 가능성이 높은 1차 선호 성향을 선택하여 기록하도록 하였다. 이후 각 전공자들의 기록지를 취합하여 도메인 별로 기록된 1차 선호 성향의 개

수를 취합하여, 분석용 도메인별 가장 많이 기록된 선호 성향을 해당 도메인의 1차 선호 성향으로 설정하였다.

이후 2차 분류에 사용될 값을 추출하는 과정을 진행하였다. 하나의 URL에 대한 2차 분류 값은 1차 분류 결과에 따라 4개의 경우의 수를 가질 수 있기 때문에 1차 선호 성향의 값 별로 하나씩 총 4개의 값을 갖는다. 각각의 값은 1차 선호유형 값이 [표 3] MBTI 유형별 주기능, 열등기능에서 주기능에 가까운 값을 갖는 유형을 2차 유형 값으로 설정하였다. 이 과정에서 동등한 우선순위의 값을 갖는 유형이 복수일 경우, E(외향)보다는 I(내향) 값을 갖는 유형에 우선 순위를 두어 선정하였다. 이후 심리학 전공자들이 2차 분류 내용을 최종 검토하는 형태로 구성하였다. 예를 들어 [표 4]를 보면, 해당 사용자의 1차 분류가 S로 확인 된 경우, 이에 종속된 ISTJ, ISFJ, ESTP, ESFP 중 해당 도메인의 1차 분석 값인 T를 주기능으로 갖는 ESTP 를 해당 도메인의 S영역 사용자 중 해당 정보를 공유할 가능성이 높은 유형이기 때문에, 2차 분류값 중 S의 값으로 설정하였다. 이와 같은 방법을 통해 2차 분류지표에 대한 DataBase를 구축하였다.

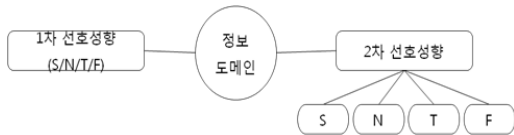


그림 2. 분석 DataBase 모형

표 2. 1차 분류 기준

분류기준	감각(Sensing)	직관(Intuition)
선호경향	인식기능 (정보수집)	
대표유형	ISTJ / ISFJ / ESTP / ESFP	INTJ / INFJ / ENTP / ENFP
분류기준	사고(Thinking)	감정(Feeling)
선호경향	의사결정 (판단)	
대표유형	ISTP / INTP / ESTJ / ENTJ	ISFP / INFP / ESFJ / ENFJ

표 3. MBTI 유형별 주기능, 열등기능

(2차 선호지표 분류 기준)

1차분류	성격유형	주기능 > 열기능
감각 (Sensing)	ISTJ	Si > Te > Fi > Ne
	ISFJ	Si > Fe > Ti > Ne
	ESTP	Se > Ti > Fe > Ni
	ESFP	Se > Fi > Te > Ni
직관 (iNtuition)	INTJ	Ni > Te > Fi > Se
	INFJ	Ni > Fe > Ti > Se
	ENTP	Ne > Ti > Fe > Si
	ENFP	Ne > Fi > Te > Si
사고 (Thinking)	ISTP	Ti > Se > Ni > Fe
	INTP	Ti > Ne > Si > Fe
	ESTJ	Te > Si > Ne > Fi
	ENTJ	Te > Ni > Se > Fi
감정 (Feeling)	ISFP	Fi > Se > Ni > Te
	INFP	Fi > Ne > Si > Te
	ESFJ	Fe > Si > Ne > Ti
	ENFJ	Fe > Ni > Se > Ti

표 4. 유입 데이터 분석 Database 예

주소	1차 분석	2차 분석값			
	S/N/T/F	S	N	T	F
http://news.naver.com/sports/	T	ESTP	ENTP	ISTP	ENFJ

### 1.4 Facebook App 개발

분석 DataBase에 대한 구성을 완료한 후, 페이스북 캔버스 앱 형태로 유형 분석 모델을 개발하였고, 페이스북 일반 사용자를 대상으로 실험을 진행하였다.

사용자가 페이스북 앱에 접근한 후, 개인정보에 대한 이용을 승인하면, 사용자의 공유 정보 중 URL 값을 가진 정보만을 추출하여 Database에 사용자 인식 ID와 함께 기록을 한 후, 1차 선호성향에 대한 분석을 진행한다. 이후 1차 선호성향에 대한 분석결과를 기록하고, 기록된 1차 선호성향에 기준하여 유입된 URL 정보들에 대한 2차 분석을 진행하여 최종 분석결과를 출력한다. 이 과정에서 공유정보 중 URL이 존재하지 않는 경우는 최종 성향으로 ISFP를 출력하도록 구성하였다. ISFP 유형은 논쟁과 갈등을 싫어하며, 자신의 의견이나 가치를 다른 사람에게 강요하지 않는 특징 때문에 정보 공유 활동에 있어서도 가장 비 활동적인 유형이기 때문이다.

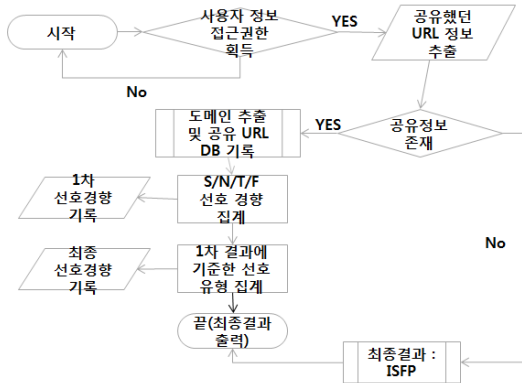


그림 3. 사용자 유형 분석 모델

## 2. 결과 분석

본 실험을 시작한 2012년 2월 16일로부터 2주의 기간 동안 약 10만 명의 순수 사용자를 확보하며 [그림 4]에서 보이는 바와 같이 빠른 속도로 확산되었으며, 2013년 10월 기준 총 200,296명의 페이스북 이용자가 본 앱을 사용하였다. 이 실험을 통해 다음과 같은 결과를 얻을 수 있었다.

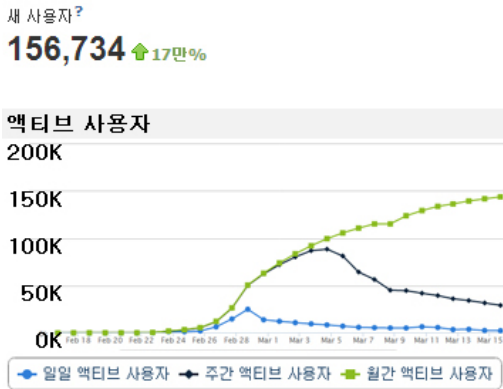


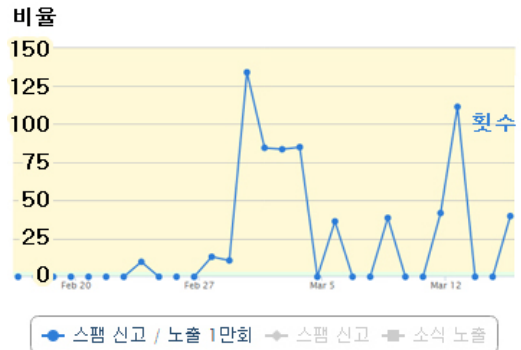
그림 4. 페이스북 오픈 후 한달 간 사용자 수 (2012/02/16 ~ 2012/03/15)

### 2.1 유형 분석 결과에 대한 사용자 반응

실험을 통한 결과에 대한 사용자 반응은 앱의 하단에 설치했던 코멘트 영역과 Facebook에서 기본적으로 제공하는 페이스북 앱 분석 서비스인 Insights의 정보 중 부정적인 피드백 영역을 통해 확인하였다.



그림 5. 실험 결과에 대한 사용자 코멘트



6.7756 스팸 신고 / 노출 1만회?      24 스팸 신고?  
35,421 소식 노출?

그림 6. Facebook 앱 Insights 정보 (부정적인 피드백)

페이스북 앱의 하단에 설치했던 코멘트 영역을 통해 분석 결과에 대한 자신의 의견을 남겨준 사용자는 총 1,356명이며, [그림 5]와 [그림 6]에서 확인할 수 있듯이 분석 결과에 대한 사용자의 긍정적인 반응은 1,215개의 코멘트로 전체 코멘트의 약 89%에 해당하는 높은 수치를 기록한 것을 확인할 수 있었다. 이는 페이스북에서 제공하는 앱에 대한 Insights 정보 중 부정적인 피드백을 통해서도 유사한 결과를 확인할 수 있는데, [그림 6]과 같이 사용자의 타임라인에 게시된 분석 결과에 대해

서도 0.00068%의 낮은 수치만이 부정적인 피드백을 주었다는 점을 확인할 수 있다.

2.2 성격유형별 분포도

본 실험에서 사용된 모델을 통해 도출된 성격 유형별 과는 [표 5]와 같다.

표 5. 본 모델을 통한 유형별 분석 결과 분포표

1차분류	성격 유형	대표 유형	인원	백분율	한국인 평균
감각 (Sensing)	ISTJ	신사임당	1,792	1.07%	25%
	ISFJ	테레사	640	0.38%	15%
	ESTP	케네디	918	0.55%	2%
	ESFP	마블린먼로	611	0.37%	2%
직관 (iNtuition)	INTJ	세종	189	0.11%	0.3%
	INFJ	간디	810	0.49%	0.3%
	ENTP	에디슨	851	0.51%	0.3%
	ENFP	체게바라	39,535	23.71%	0.3%
사고 (Thinking)	ISTP	이소룡	370	0.22%	5%
	INTP	아인슈타인	257	0.15%	1%
	ESTJ	라이스	736	0.44%	15%
	ENTJ	나폴레옹	705	0.42%	0.3%
감정 (Feeling)	ISFP	베토벤	107,289	64.35%	7%
	INFP	앤디워홀	385	0.23%	1%
	ESFJ	루피	9,100	5.46%	7%
	ENFJ	링컨	2,534	1.52%	0.3%
총합			166,722	100%	100%

[표 5]와 [그림 7]에서 보는 바와 같이 각 유형별 분포 중 2개의 특정 성격 유형에 대해 밀집된 결과가 발생한 것을 볼 수 있는 데, 외향적이며 상상력이 풍부한 ENFP 유형에 대한 밀집 현상은 Youtube의 동영상 콘텐츠가 페이스북 상에서 공유된 다른 사이트와는 다르게 기형적으로 5만회 이상 공유가 되었지만 분류 과정에서 이에 대한 가중치를 반영하지 않았기 때문에 추측된다. ISFP 유형은 사용자의 공유 정보의 수가 기준 양에 못 미칠 경우 혹은 분석된 URL이 없을 경우에 대한 대표 유형으로 설정하였었다. 본 실험 결과 64.35%의 사용자가 ISFP 유형으로 분류되었는데, 이는 한국인의 경우 내향적인 성향이 64%에 이른다는 심해숙의 연구 결과[15]와 유사하다. 이러한 사항은 내향적 성향

이 페이스북의 정보 공유에 있어서도 비극적성을 띄는 것으로 추측되며, 향후 연구에서 이에 대한 검증 및 보안 작업을 통해 성격유형 분석의 타당성을 높일 수 있을 것이라 판단된다.

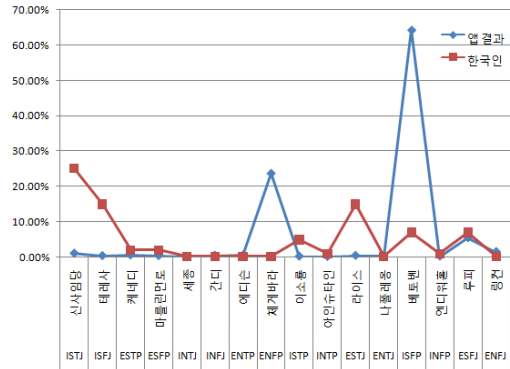


그림 7. 유형별 분포도 비교도

[그림 7]의 비교도를 통해 한국인의 MBTI 유형분포 [15]와 비교해보면, 일부 유형에서는 차이를 보이긴 하지만, 10개의 유형의 분포도는 비슷한 형태를 보이고 있음을 확인할 수 있다.

2.3 정보 제공 사이트별 공유 횟수 분석

본 실험을 통해 수집된 URL에서 특이점을 찾기 위해 추가 분석을 진행하였다. 본 실험 참가자들은 페이스북 상에서 총 75,348 개의 사이트에서 만들어진 1,846,360 개의 URL정보를 공유하고 있음을 확인할 수 있었다.

표 6. URL 공유 횟수 구간 별 사이트 수

URL 공유 횟수	사이트 수
50 만회 이상	1
10만회 이상 50만회 미만	0
1만회 이상 10만회 미만	13
1천회 이상 1만회 미만	117
100회 이상 1천회 미만	673
10회 이상 100회 미만	4,223
10회 미만	70,321

이 중 성격유형 분석 결과의 정확성을 높이기 위해

[표 6]과 같이 URL 공유 횟수별 사이트 수를 구분하였고, 1만회 이상 공유되는 14개의 사이트에 종속된 하위 URL에 대해서는 가중치를 줄일 필요가 있기 때문에, 이 사이트들에 대한 정보를 [표 7]와 같이 추출하였다.

표 7. 1만회 이상 공유된 사이트

1만회 이상 공유된 사이트	count	성격	백분율
www.youtube.com	728,923	영상	39.48%
krhoroscopes.wantforall.com	151,791	편	8.22%
www.facebook.com	26,125	페이스북	3.36%
apps.facebook.com	17,291	페이스북 앱	0.94%
news.nate.com	16,691	뉴스	0.90%
www.hani.co.kr	16,343	뉴스	0.89%
blog.naver.com	15,828	블로그	0.86%
vimeo.com	15,109	영상	0.82%
news.naver.com	14,845	뉴스	0.80%
www.ohmynews.com	14,008	뉴스	0.76%
media.daum.net	13,768	뉴스	0.75%
news.chosun.com	13,209	뉴스	0.72%
tvpot.daum.net	13,186	영상	0.71%
www.review.kr	12,464	편	0.68%

이를 통해 페이스북 상에서 공유되는 정보들 중 영상 콘텐츠를 제공하는 사이트의 수는 적지만, 공유횟수로는 54%의 많은 횟수가 이뤄지고 있음을 확인할 수 있으며, 편 영역의 퍼센트를 합산해보면, 뉴스 같은 정보성 콘텐츠 보다는 재미, 흥미에 기반을 둔 엔터테인먼트 적인 콘텐츠의 소비가 약 69%로 대부분을 차지하고 있음을 알 수 있다.

### 3. 결론 및 향후 연구 과제

현재 소셜 미디어를 통해 누적되고 있는 사용자 데이터는 그 활용에 따라 개별의 사용자의 특징을 파악하여 개인/집단에 최적화된 서비스를 제공할 수 있으며, 해당 콘텐츠가 사용자에게 주는 공감의 정도에 따라 사용자가 보다 적극적이고 능동적으로 참여하는 환경이 만들어졌음을 확인할 수 있었다. 이러한 새로운 환경에서 소셜 미디어의 데이터 활용 및 분석에 관한 연구는 다양한 분야가 서로 융합되어 연구와 실험이 진행될 필요가 있다.

본 실험을 통해 사용자가 소셜 미디어 상 공유한 정보는 단순히 온라인 상에서 소비하는 정보들과는 다르게 사용자의 기호가 반영되었으며, 이를 통해 사용자의 성격 유형을 파악할 수 있는 기초 단서가 될 수 있음을 확인할 수 있었다. 향후 연구에서는 분석 Data Base의 범위를 더욱 늘려 예외로 처리되는 숫자를 줄이고, 1만회 이상 공유된 사이트에 대해 가중치를 주는 연구를 진행함으로써 도출 결과에 대한 완성도를 높이는 것이 필요할 것으로 보인다.

### 참고 문헌

- [1] <http://www.statista.com/topics/840/smartphones/chart/1405/the-united-states-ranks-13th-in-smartphone-penetration/>
- [2] [http://allfacebook.com/2q-2013\\_b122375](http://allfacebook.com/2q-2013_b122375)
- [3] [http://magazine.hankyung.com/jobnjoy/apps/news?popup=0&nid=05&c1=5002&nkey=2013071500039049265&mode=sub\\_view](http://magazine.hankyung.com/jobnjoy/apps/news?popup=0&nid=05&c1=5002&nkey=2013071500039049265&mode=sub_view)
- [4] <http://ko.wikipedia.org/wiki/소셜미디어>
- [5] 김태우, “SNS의 스마트폰 게시글 내용 분석을 통한 사용자의 요구특성 분석”, 한국전자거래학회지, 제17권, 제4호, pp.197-208, 2012.
- [6] 강부식, “사회연결망의 구조적 공백을 활용한 신규고객 웹 상품추천방법”, 산업경제연구, 제23권, 제5호, pp.2371-2385, 2010.
- [7] 옥경영, “관계네트워크의 정보 확산에서 혁신자와 허브가 Early Takeoff와 확산크기에 미치는 영향 연구”, 한국콘텐츠학회논문지, 제12권, 제9호, pp.348-356, 2012.
- [8] 박한우, “월드와이드웹에 나타난 국제 학술 커뮤니티 네트워크에 대한 탐사적 연구”, 한국문헌정보학회지, 제37권, 제2호, pp.153-168, 2003.
- [9] <http://ko.wikipedia.org/wiki/빅데이터>
- [10] [http://ko.wikipedia.org/wiki/데이터\\_마이닝](http://ko.wikipedia.org/wiki/데이터_마이닝)
- [11] <http://blog.naver.com/PostView.nhn?blogId=someonenice&dlogNo=10045558029>
- [12] <http://ko.wikipedia.org/wiki/mbti>



- [13] 오시룡, 석혜정, "3D 애니메이션의 캐릭터 유형 및 성격 분석 :픽사의 애니메이션을 중심으로", 만화애니메이션연구, 통권 제8호, pp.161-183, 2005.
- [14] 황은일, 김택호, 김각숙, "MBTI 선호지표에 따른 만다라그림의 표현양식", 미술치료연구, 제20권, 제1호, pp.67-84, 2013.
- [15] 김정택, 심혜숙, "성격유형검사(MBTI)의 한국 표준화에 관한 일연구", 한국심리학회지: 상담 및 심리치료, 제3권, 제1호, pp.44-72, 1990.

저 자 소 개

한 정 화(Jung Hwa Han)

준회원



- 2007년 2월 : 서울 디지털 대학교(디자인 학사)
- 2007년 3월 ~ 현재 : 중앙대학교 첨단영상대학원(예술공학 석사과정)
- 2011년 6월 ~ 현재 : 트리플 대표

<관심분야> : 소셜 데이터 활용, 적정기술

박 진 완(Jin Wan Park)

정회원



- 1995년 2월 : 중앙대학교 컴퓨터 공학과(공학사)
- 1998년 : Pratt CGIM Computer Media(MFA)
- 2003년 3월 ~ 현재 : 중앙대학교 첨단영상대학원 교수

<관심분야> : Art&Technology, Procedural Animation