

스토리데이터 설계, 구축 및 창작 요소 분포 분석

김지선¹ · 김명준^{2*}¹이화여자대학교 융합콘텐츠학과 석박사통합과정^{2*}이화여자대학교 융합콘텐츠학과 교수

Analysis of the Distribution of Creative Elements in StoryData

Ji-Seon Kim¹ · Myoung-Jun Kim^{2*}¹Master-Doctoral Combined Course, Dept of Content Convergence, Ewha Womans University, Seoul 03760, Korea^{2*}Professor, Dept of Content Convergence, Ewha Womans University, Seoul 03760, Korea

[요약]

본 논문은 대중문화 스토리 콘텐츠를 분석한 스토리데이터 구축한 사례에 대해 소개하고, 스토리 데이터에 나타난 창작 요소의 분포를 분석한다. 스토리 콘텐츠를 작품 단위와 서사의 최소 단위인 유닛(unit) 단위로 분석하여 구축되었으며 장르, 주제, 모티프, 서사단계 등의 서사 요소를 포함한 상세한 메타데이터를 포함하고 있다. ‘Sudden Accident’, ‘Love Triangle’, ‘Unrequited Love’ 등의 모티프가 높은 빈도를 보였고 대중문화 콘텐츠의 스토리가 사랑과 인간관계의 소재를 중심으로 드라마와 로맨스 장르를 포괄하는 경향을 보인다는 사실을 정량적으로 확인하였다. 또한 대중문화 콘텐츠 분석 시 유닛에 대한 분석 및 재창작 과정에서 감정 형용사와 행위 동사를 중심으로 메타데이터를 추가하여 대중적인 스토리의 생동감과 감정적 요인을 포함하는 분석 및 재창작을 거쳐 데이터베이스를 구축하였다. 또한 고전 텍스트에 초점을 맞춘 기존 데이터베이스와는 달리, 이 연구는 현대 대중문화의 내러티브 구성 요소와 창의적 요소를 정량화하여 새로운 관점을 제시하고 스토리 관련 인공지능 도구의 토대를 제공한다.

[Abstract]

This paper presents the construction of StoryData, a specialized database to underpin the development of an AI-based story generation model and creative support service using narratives from contemporary media such as films, dramas, web novels, and comics. This database constructs and reconstitutes popular culture content at both the work and unit level - the smallest narrative unit - incorporating detailed metadata that encompasses genres, themes, motifs, and narrative stages. A quantitative analysis reveals motifs such as “Sudden Accident,” “Love Triangle,” and “Unrequited Love” are prevalent, indicating a trend in popular content that gravitates toward drama and romance with a focus on love and human relationships. The research also enriches the database with metadata focusing on emotional adjectives and action verbs during the analysis and re-creation of units, capturing the vibrancy and emotional aspects of popular stories. In addition, this research quantifies the narrative and creative elements of contemporary popular culture, offering a novel perspective and laying the groundwork for AI tools that aim to expedite and enhance the storytelling process.

색인어 : 스토리 데이터, 인공지능, 창작 요소, 데이터베이스, 내러티브 분석**Keyword** : StoryData, Artificial Intelligence, Creative Elements, Database, Narrative Analysis<http://dx.doi.org/10.9728/dcs.2023.24.12.3005>

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 13 November 2023; Revised 07 December 2023

Accepted 14 December 2023

***Corresponding Author; Myoung-Jun Kim**

Tel: +82-2-3277-3347

E-mail: mjkim@ewha.ac.kr

1. 서론

1-1 연구배경 및 목적

작가의 창작 행위에 도움을 주기 위하여, 작가의 구상에 따라 스토리 추천 및 자동 생성을 위한 인공지능 모델 개발을 위해서는 스토리의 데이터베이스가 필요하다. 이화여자대학교 연구진은 스토리 시놉시스 아이디어이션 데이터베이스 <스토리헬퍼>를 구축 및 서비스하였다[1]-[6]. 최미란[7]은 고소설을 캐릭터와 플롯으로 나누어 데이터베이스화하는 모델을 제안하였다. 전대진[8]은 서사 이론을 고려하여 설화 구조, 전형적인 캐릭터, 공간, 사건 동기를 기준으로 설화를 분석하여 수용자가 쉽게 사용할 수 있는 데이터베이스를 구축하고자 하였다. 김은정, 신동일, 황수경[9]은 영문학 작품에서 모티프를 추출, 분석하여 DB화하여 창작 지원 시스템을 만들고자 하였다. 이재열[10]은 구비설화를 데이터베이스화하는 과정에서 기존의 ‘서사물’ 데이터베이스가 단순히 텍스트를 디지털화하고 메타 데이터를 부가하는 정도에 그친다는 한계를 지적하며, ‘서사물’을 다루는 데이터베이스는 서사 자체에 대한 정보를 포함하며 이를 통해 서사의 정보를 조작할 수 있어야 한다고 밝힌다. 이상호[11]는 문화 콘텐츠 개발의 원천으로서 한국고전번역원이 구축한 한문고전 DB 및 활용 콘텐츠를 분석하였다. 이현수, 김정미[12]은 고전 서사 콘텐츠가 아닌 대중문화 콘텐츠로서 스토리 데이터베이스 구축까지 도달하지는 못하였으나 캐릭터 웹드라마의 기승전결에 따른 이야기 틀을 분석하고 클리셰를 목록화하여 정리한 후 GPT를 활용한 스토리 추천 모델을 개발하였다.

본 논문에서는 영화, 드라마, 웹소설, 웹툰/만화의 대중문화 서사물 콘텐츠의 스토리를 분석 및 재창작하여 서사적 정보를 포함하는 스토리 데이터베이스를 구축하고, 스토리 데이터에 나타난 창작 요소의 분포를 분석하여 대중문화 콘텐츠에 나타나는 스토리에 대한 이해를 제공하고자 하였다.

1-2 스토리데이터 설계 및 구축

본 연구진은 인공지능 학습을 위한 데이터 구축의 일환으로[13] 영화, 드라마, 웹소설, 만화/웹툰을 대상으로 대중문화 스토리 콘텐츠를 분석하여 2022년 6월 1일부터 2022년 12월 31일까지 10만여건의 스토리데이터를 구축하였다. 본 논문에서는 데이터 설계 및 구축 과정에 대해 소개하고 데이터에 나타난 창작 요소에 대해 분석하고자 한다. 각 작품이 유동적인 범위 내에서 이론적인 서사 단계를 따른다고 가정하고 분석을 진행하였고, 작품에 대한 서사 정보를 메타데이터로 분석했다. 먼저 작품 전체 단위의 창작요소를 분석하여 데이터로 작성하였고, 스토리를 분석단위인 여러개의 유닛(unit)으로 구분하고, 유닛 별로 내용을 재창작하고 및 창작요소를 분석하여 데이터를 작성하였다. 이때, 캐릭터는 비식

표 1. 스토리데이터 데이터 예시

Table 1. Example of StoryData

Title	To go Home
Genre	Drama
Theme	Love
Concept	A tired office worker, C001, who is exhausted from daily life, meets the ever-joyful C002 and discovers the joys of life
Structure	15 Stage of Storyhelper
Motif	Love at First Sight

Unit 1, 2, . . .

Character	C002, C004, C005, C012
Stage	Setting-up
Storyline	C004 scolds the unemployed C002, after which C002 reluctantly gets up and goes to look for a job.
Causality	First Action

Location	Character	Act	Emotion	Content
Home	C004	Go	Angry	Angry, they go to C004's room and fling open the door
Home	C004	Yell	Angry	How long are you going to sleep!
Home	C002	Sleep		C002 ignores it and continues to sleep
Home	C004		Angry	Seeing C002's calm demeanor makes her even angrier
...

별화하였다. 분석한 스토리데이터는 표 1과 같다.

작품 단위에서는 유형, 장르, 주제, 기본 설정, 서사구조의 종류, 모티프, 주인공 유형, 작품의 주요 갈등구조 등을 창작요소로 보고 분석하여 데이터를 작성하였다. 장르는 한 작품당 최소 1개에서 최대 3개로 하여 수집하였다. 스토리 데이터 전반적으로 장르별 작품수를 집계하는 데 있어, 장르가 2개 이상인 작품에 대해서는 일정한 비율을 적용하였다. <스토리헬퍼>[3]에서는 스토리 분석을 수행을 위해 이재선이 서사 주제학에서 제안한 모티프 항목과 이승진이 독일문학 모티프 DB 구축 시 제안한 모티프 항목을 종합하여 총 204개 모티프를 사용하였다. 본 연구에서는, 최근의 창작성향을 고려하여 웹소설 및 웹툰에서 많이 활용되는 모티프 3개 항목을 추가하였다. 스토리데이터는 표 2의 총 207개의 모티프 항목을 적용하여 분석을 진행하였다[3],[14],[15].

작품의 스토리 유닛 단위 분석은 유닛의 메타데이터와 스토리데이터로 구성하였다. 유닛의 스토리데이터는 인물의 행위 또는 생각, 상황 혹은 분위기를 묘사하는 ‘Narrative’ 문장과 해당 맥락에서 인물의 대사를 간략히 드러내는 ‘Script’로 문장으로 구성하였다. 각 문장은 행위 인물, 행위 동사, 감정 메타데이터를 추가하여 분석하였다. 유닛의 전반적인 메타데이터 분석은 다음과 같이 수행하였다. 해당 유닛이 작품에서 서사적

표 2. 스토리데이터 작품 모티프 항목

Table 2. Motif items of StoryData

Category	Motif Items
Situation	Obsessive-Compulsive Disorder, Alternate History, Insomnia, Biological Catastrophe, Pact with the Devil, Post-Traumatic Stress Disorder, Midlife Crisis, Reincarnation, Borderline Personality Disorder, Urban Disaster, Asymmetric Warfare, Gender Identity Confusion, Alcoholism, Brave Encounter, Apocalypse, Dimensional Travel, Panic Disorder, Apprenticeship, Possession, Terminal Illness Diagnosis, Loss of a Lover, Depression, Collective Hysteria, Challenge of Old Age, Defamation of Honor, Accidental Death, Colonization, Prophecy, Artificial Hibernation, Natural Catastrophe, False Accusation, Castaway on a Desert Island, Death of a Loved One, Child Abuse, Alien Invasion, Outbreak of War, Juvenile Delinquency
Character	Matriarch of the Family, Arrival of a Savior, Seducer for Money, Womanizer, Elderly in Love, Mischievous Person, Disabled Person, Exorcist, Sudden Illness, Summoning Spirits, Mutation, Half-human, Half-animal, Cyborg, Devil's Child, Visitor from the Underworld, Bankruptcy, Robbery, Amnesia, Homosexual, Traitor, Psychopath, Diabolical Artist, Prisoner of War, Pygmalion Effect, Giant Monster, Windfall, Intelligent Fool, Vampire, Contract Killer, Femme Fatale, Zombie, Abusive Stepfather, Noble Prostitute, Eccentric Teacher, Unexpected Letter, Vampire Hunter, Young Hero, Superior Subordinate, Disappearance of an Acquaintance, Pirate, Noble Maid, Multiple Personalities, Sorcerer, Disguise, Martyr, Rebellion Against Fate, Prostitute, Homo Superior, Tormented Gangster, Thief, Drug Dealer, Abandoned Child, Hidden Lover, Wanderer, Virgin Hero, Sacrificing Parent, Con Artist, Inept Hero, Decline of Nobility, Clone, Marriage for Social Climbing, Righteous Outlaw, Angel, Sacrificing Child, Loan Shark, Doppelgänger, Unknown Guardian, Corrupt Police, Identity Disguise, Hostage, Careerism, Obsessive Stalker, Noble Savage, Dictatorship, Zealous Evangelist, Unfortunate Genius, Outsider, Autism, Dementia, Northern Archduke
Action	Family Division, Accusation, Relocation to an Unfamiliar Place, Lolita Complex, Spousal Dispute, Wife Swapping, Oedipus Complex, Electra Complex, Unrequited Love, Adultery, Conflict with Mother-in-law, Necrophilia, Masochism, Unwelcome Guest, Stalking, Accidental Murder, Suicide, First Sexual Experience, Imprisonment, Voyeurism, Major Combat, Rebellion, Sadism, Time Travel, One-Night Stand, Curse, Love at First Sight, Search for Hidden Lineage, Mad Love, Massacre, Armed Uprising, Doll That Became Human, Cyrano Complex, Dangerous Secret Marriage, Cursed Place, Pursuer, Sudden Accident, Marriage Proposal Challenge, Great Voyages and Expeditions, Unsolved Murder Case, Pretense of Love, Wife Stealing, Ghost, Love with the Enemy, Rediscovery of Preferences, Rape, Homecoming, Fugitive, Beautiful Simpleton, Love Triangle, Child Abduction, Othello Syndrome, Friendship with the Enemy Army, Escape from Prison, Forced Task, Incest, Gambling, Treasure Hunt, Sexual Dysfunction, Beautiful Affair, Instigation, Rescue of a Comrade, Terrorism, False Statement, Falling for the Bad Guy, Friendship with Animals, Vendetta, Friendship Across Generations, Serial Murder, Visit to Another World, Confession of Sin, Defrocking, Love Across Class Lines, Love Overcoming Age Difference, Joint Suicide, Unjust Discrimination, Investigation Request, Appearance of an Ex-Lover, Bullying, Geographic Isolation, Strike, Isolated Couple, Kidnapping, Sudden Superpowers, Search for Parents, Riddle/Enigma, Misunderstanding, Divorce, Mocking the Ruler, Sibling Rivalry

위치를 나타내는 서사 단계, 스토리라인 요약, 해당 유닛과 맥락적으로 연결된 다른 유닛 번호, 유닛 단위에 적용 가능한 모티프, 인과관계(Causality)을 창작 요소로 보고 분석을 진행했다. 행위 동사와 감정 형용사는 모두 한국어의 표현 범주에 대한 연구 내용을 고려하여 목록을 생성하였다[16],[17].

II. 스토리데이터 창작 요소 분포 분석

구축된 스토리데이터에 나타난 창작 요소의 빈도를 분석하여 대중문화 스토리텔링 콘텐츠의 특징을 포괄적으로 파악하고자 한다. 스토리데이터의 장르, 주제, 모티프, 감정 형용사, 행위 동사, 유닛 인과성(Causality), 서사단계를 주요 창작 요소로 보고 이에 대한 빈도를 분석하였다.

2-1 스토리데이터의 장르 빈도 분석

그림 1은 스토리데이터에서 각 장르의 비율을 나타낸다. 구축된 스토리데이터의 장르는 그림 1과 같이 나타났다. Drama(39%), Romance(24%)로 두 장르의 총합이 전체 스토리데이터의 반 이상을 차지했으며, 이어서 Thriller(13%),

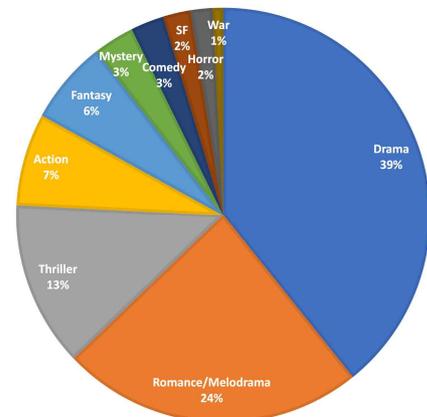


그림 1. 스토리데이터의 장르 분포 분석

Fig. 1. Proportional representation of genres in StoryData

Action(7%), Fantasy(6%) 등이 상위 5개 장르로 나타났다. Mystery, Comedy, SF, Horror, War 등은 각 5% 이하의 적은 비율로 나타났다. Drama와 Romance 장르의 비율이 과반을 차지하며 장르 다양성 측면에서 특정 장르가 우세하게 나타난다는 것을 확인할 수 있다. 이때, Drama, Romance, Melodrama 장르의 비율로 미루어 보아의 감정과 관계를 중심으로 한 스토리가 대중문화 서사 콘텐츠에서 중

요한 부분을 차지하며 감정과 인간 관계에 대한 깊은 탐구가 대중문화 콘텐츠 소비자에게 지속적으로 강한 호소력을 발휘하고 있다고 추측할 수 있다.

2-2 스토리데이터의 주제 빈도 분석

그림 2는 구축된 스토리데이터의 주제의 빈도를 나타낸다. Love의 비율이 전체의 30%를 차지하며, 이어지는 주제인 Challenge(10%), Maturation(6%), Adventure(5%), Pursuit(5%) 와도 큰 차이로 높은 비율을 차지하는 것으로 나타났다.

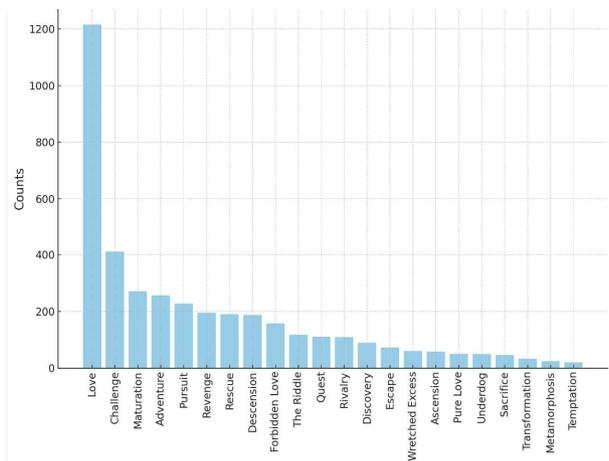


그림 2. 스토리데이터의 주제 분포 분석
Fig. 2. Distribution of themes in StoryData

대중문화 콘텐츠에서 사랑이 매우 보편적인 주제이며 사랑이 인간의 보편적 경험 및 감정을 반영하는 주제로서 대중문화 콘텐츠에서 중요한 요소로 자리 잡았음을 시사한다. 사랑은 대중문화 콘텐츠에서 스토리를 추동하는 동기 부여 요소로써, 갈등, 성장, 해결 등의 요소와 결합하여 대중적인 공감을 이끌어내고 있다고 추측할 수 있다. 또한 도전과 성숙의 주제가 사랑에 뒤이어 높은 비율을 차지하는 것으로 보아 대중문화 콘텐츠에서 개인의 성장 및 발전 역시 중요한 주제로 다루어지고 있음을 확인할 수 있다. 또한 대중문화 콘텐츠 주제로서의 모험과 추구의 분포로 미루어 보아 모험 추구와 목표 달성을 위해 노력하는 스토리가 대중문화 콘텐츠에서 긴장감과 흥미를 유발하며 이야기에 동력을 부여하고 관객을 몰입시키는 요소로 작용하고 있다고 추측할 수 있다.

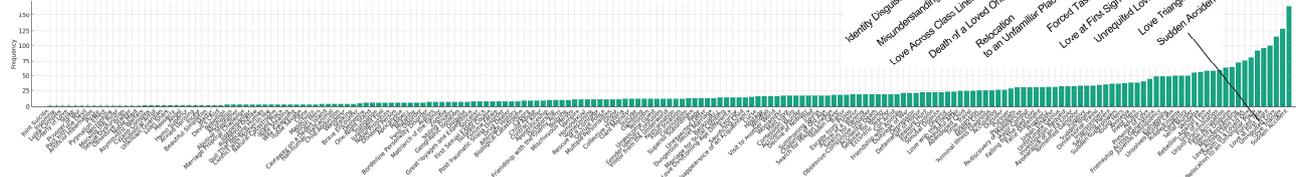


그림 3. 스토리데이터의 모티프의 빈도 분석
Fig. 3. Frequency distribution of narrative motifs in StoryData

2-3 스토리데이터 모티프 빈도 분석

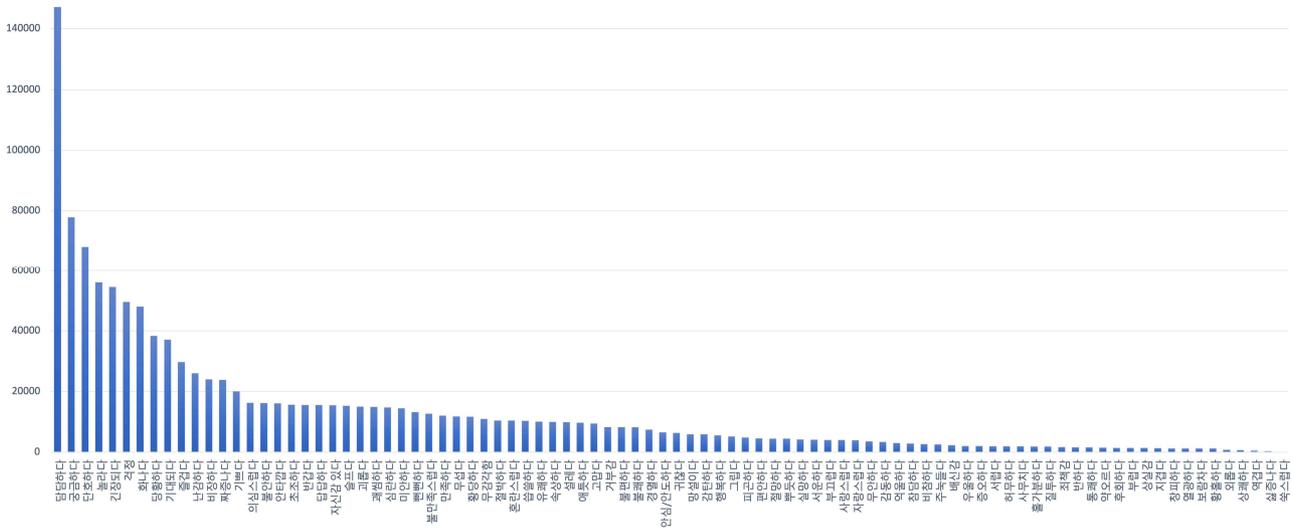
그림 3은 대중문화 서사 콘텐츠를 바탕으로 구축된 스토리데이터에서 나타난 서사적 모티프의 빈도이다. 빈번한 모티프들은 전체 빈도의 상당 부분을 차지하는 반면, 덜 흔한 모티프들은 전체 분포에서 비교적 작은 비중을 차지한다. 빈도가 높은 모티프에서 매우 낮은 모티프까지 분포가 고르면서 빈도 높은 것이 더욱 높은 이차함수의 모양을 보이고 있다. 이는 대중문화 콘텐츠에서 상대적으로 보편적인 모티프가 더욱 보편적으로 사용되어 빈도가 더욱 높은 경향이 있음을 시사한다.

상위 10개 항목은 ‘Sudden Accident’, ‘Love Triangle’, ‘Unrequited Love’, ‘Love at First Sight’, ‘Forced Task’, ‘Relocation to an Unfamiliar Place’, ‘Death of a Loved One’, ‘Love Across Class Lines’, ‘Misunderstanding’, ‘Identity Disguise’이며, 각 항목이 총합 전체 스토리데이터의 25%를 차지한다. 그림 3에서 가장 흔한 모티프는 “Sudden Accident”이며, 이는 대중문화 콘텐츠에서 갑작스러운 사건이 중요한 전환점을 제공하는 주요 요소임을 나타낸다. ‘Love Triangle’, ‘Unrequited Love’, ‘Love at First Sight’와 같은 사랑과 관련된 모티프들이 모두 높은 빈도로 나타나는 것으로 보아, 사랑과 관계는 대중문화 콘텐츠에서 중요한 주제임을 확인할 수 있다.

2-4 스토리데이터의 감정 형용사 빈도 분석

그림 4는 스토리데이터에 나타난 감정 형용사의 빈도를 나타낸다. 스토리데이터에서 감정 형용사는 총 1,195,917회 나타났다. 그 중 ‘답답하다’가 146,938회로 가장 많이 나타나 전체 비율의 12.99%를 차지한다. 상위 5가지 감정 형용사는 ‘답답하다’, ‘궁금하다’, ‘단호하다’, ‘놀라다’, ‘긴장되다’로 전체 감정 형용사 빈도의 33.71%를 차지한다. 가장 적게 나타난 감정 형용사의 빈도는 ‘쑥스럽다’ 1회이다. 즉, 감정 형용사는 평균적으로 큰 차이를 보이며 높은 변동성이 있음을 의미한다.

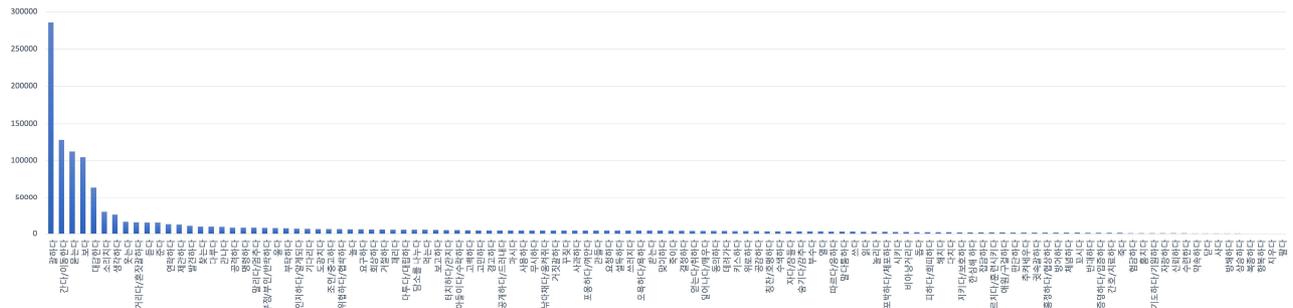
상위 5 항목의 감정 형용사가 전체의 삼할을 차지하는 것으로 보아, 대중문화 콘텐츠에서는 특정 감정들이 중심적인 역할을 하고 있다. 이를 통해 스토리 창작 시 특정 감정적 요소를 강조함으로써 소비자들에게 강렬한 인상을 남기기 위한 전략이 사용되었을 것이라고 추측할 수 있다. 또한 가장 빈번



*The items are presented in Korean to show the actual data.

그림 4. 스토리데이터의 감정 형용사 분포

Fig. 4. Distribution of emotional descriptors in StoryData



*The items are presented in Korean to show the actual data.

그림 5. 스토리데이터의 행위 동사 분포

Fig. 5. Distribution of action descriptor in StoryData

한 감정이 ‘답답하다’라는 것은, 대중문화 콘텐츠에서 침착하거나 내면적인 순간이 중요하게 작용하고 있음을 드러낸다. 주인공이나 중요 캐릭터가 상황을 잘 통제하고 있거나, 감정을 내면화하는 경향이 많이 표현되고 있음을 나타낸다.

2-5 스토리데이터의 행위 동사 빈도 분석

그림 5는 스토리데이터에 나타난 행위 동사의 빈도를 나타낸다. 스토리데이터에 행위 동사는 총 1,277,624회 나타났으며 그 중 ‘말하다’가 286,097회로 가장 많이 나타나 전체 비율의 22.39%를 차지한다. 상위 5가지 행위 동사는 ‘말하다’, ‘간다/이동한다’, ‘묻는다’, ‘보다’, ‘대답한다’로 전체의 54.37%를 차지한다. 즉, 행위 동사의 분포는 감정 형용사와 마찬가지로 평균적으로 큰 차이를 보이며 높은 변동성이 있음을 의미한다.

‘말하다’, ‘묻는다’, ‘대답한다’와 같은 언어적 행동을 의미하는 동사가 높은 빈도로 나타나는 것으로 미루어 보아 대중문화 콘텐츠에서 대화가 캐릭터의 개성, 갈등, 관계 발전을 풍

부하게 하는 중심적인 역할을 한다고 추측할 수 있다. 또한 ‘간다/이동한다’와 같은 물리적 행동의 빈도가 두 번째로 높은 것으로 미루어 보아 이동이 대중문화 콘텐츠에서 마찬가지로 중요한 역할을 하며, 캐릭터가 새로운 환경에 직면하거나 단순히 이동하면서 스토리가 전개된다는 것을 의미한다. 또한 ‘간다/이동한다’와 같은 이동을 나타내는 행위 동사와 ‘보다’와 같은 시각적 인식을 나타내는 행위 동사가 높은 빈도를 보이는 것으로 미루어 보아, 대중문화 콘텐츠에서는 인물의 물리적이고 관찰적인 역동성이 두드러지며 인물이 인물 또는 환경과 상호작용하는 것이 서사 진행에 있어 중요하게 작용한다고 추측할 수 있다.

2-6 스토리데이터의 유닛 인과성 빈도 분석

그림 6은 스토리데이터 유닛 단위 분석에 나타난 인과성 데이터의 빈도를 나타낸다. 스토리데이터에 총 유닛 인과성은 86,229회 나타났으며 그 중 ‘Ally’가 5,597회로 가장 많이 나타나 전체 비율의 6.49%를 차지한다. 상위 10가지 유닛 인

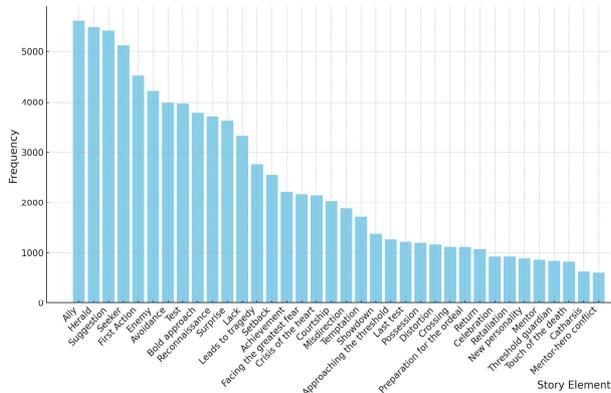


그림 6. 스토리데이터의 유닛 인과성 분포
Fig. 6. Distribution of causalities in popular culture narratives

과성 항목은 ‘Aim’, ‘Herald’, ‘Suggestion’, ‘Seeker’, ‘First Action’, ‘Enemy’, ‘Avoidance’, ‘Test’, ‘Bold Approach’, ‘Reconnaissance’로 전체의 30.28%를 차지한다. ‘Mentor-hero conflict’가 611회로 가장 적은 빈도로 나타났다. 스토리데이터의 유닛 인과성은 다양한 분포를 보이지만 특정 중요한 인과성 항목이 강조 및 집중되는 경향이 있음을 나타낸다. 상위 항목 중 ‘Ally’는 캐릭터 간의 연대와 협력을, ‘Herald’는 스토리에서 변화의 시작 혹은 중요한 사건의 전령으로 기능하는 캐릭터나 사건을 의미한다. 이는 인물 간 연대와 협력이 대중문화 콘텐츠에서 소비자의 공감을 이끌어내는 중요한 요소이며 서사 전개 과정에서 사건의 추동 동기 부여가 중요하게 작용한다는 것을 추측할 수 있다. 또한 ‘Seeker’나 ‘First Action’의 높은 빈도는 대중문화 콘텐츠에서 캐릭터의 주도적인 행동의 중요성을 드러낸다. 즉, 이러한 인과성 항목들은 주인공이 목표나 목적을 가지고 모험의 부름(전령)을 받고 진행 방법에 대한 힌트나 제안을 받는 많은 스토리의 초기 단계를 나타낸다. 즉, 대중서사 스토리 콘텐츠에서 내러티브의 기반 및 방향을 설정하는 것에 분량을 할애하고 있으며, 캐릭터가 자신의 목표를 적극적으로 추구하고 문제를 해결해 나가는 과정을 중심으로 전개되어 간다는 사실을 반영한다.

2-7 스토리데이터의 서사 단계 빈도 분석

그림 7은 스토리데이터에 나타난 유닛 별 서사단계의 빈도를 나타낸다. 서사단계는 <스토리헬퍼>의 데이터 분석에서 도입 하였던 3막 구조 15단계를 기반으로 분석하였다. ‘Setting-up’이 총 9,315회로 가장 높은 빈도로 나타났으며 ‘Trailer Moments’와 ‘Final Salvo’가 각각 가장 적은 빈도로 나타났다. 스토리헬퍼 15단계를 기반으로 한 스토리데이터의 경우 1막에 해당하는 유닛이 총 31,894회, 2막에 해당하는 유닛이 총 37,927회, 3막에 해당하는 유닛이 총 20,327회 나타났다.

전형적인 영웅서사 작품의 경우는 유닛의 조셉 캠벨의 영웅의 12단계를 기반으로 서사단계 분석을 하였다. 그림 8은 스토리데이터에 나타난 영웅의 12단계를 기반으로 분석한 데이터의 서사단계 별 빈도를 나타낸다. ‘Tests, Allies, and

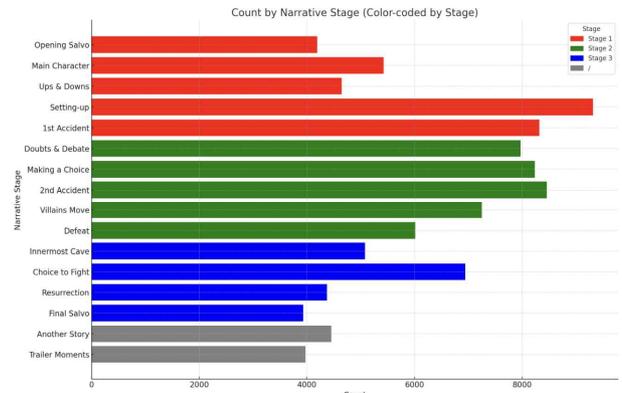


그림 7. 스토리데이터의 유닛 서사단계 분포(스토리헬퍼 15단계 기준)

Fig. 7. Distribution of unit narrative stages in StoryData (based on the 15 stages of StoryHelper)

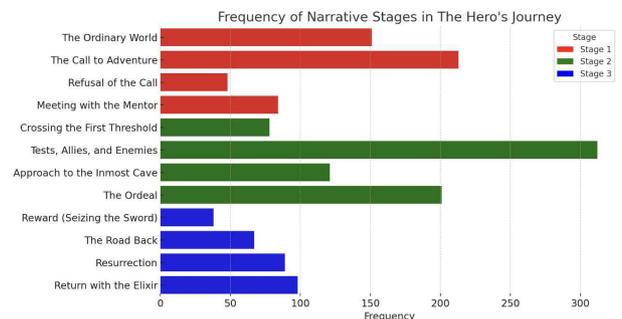


그림 8. 스토리데이터의 유닛 서사단계 분포 (영웅 12단계 기준)
Fig. 8. Distribution of unit narrative stages in StoryData (based on the hero's 12 stages)

Enemies’가 312회로 가장 많이 나타났으며 ‘Reward’가 38회로 가장 적게 나타났다. 영웅의 12단계를 기반으로 한 스토리데이터의 경우 1막에 해당하는 유닛이 총 574회, 2막에 해당하는 유닛이 총 672회, 3막에 해당하는 유닛이 총 254회 나타났다.

그림 7과 그림 8 모두에서 스토리의 갈등이 본격적으로 진행되는 2막의 총 빈도수가 가장 높게 나타났으며, 캐릭터와 상황에 대한 설명이 이루어지는 1막이 갈등이 해소되는 3막보다 더 높은 빈도로 나타났다. 이는 대부분의 대중문화 콘텐츠 스토리는 주인공을 중심으로 전개되며, 주인공이 시련을 겪고 동맹이나 적을 만나 성장하는 과정이 핵심적인 서사 요소가 되고 있으며 수용자가 캐릭터에게 공감할 수 있는 기반을 제공하고 갈등의 주된 배경에 대한 이해도를 제고하는 것이 중요함을 의미한다.

III. 결 론

본 논문에서는 본 연구진이 수행한 대중문화 콘텐츠의 스토리 데이터 구축 사례에 대해 소개하고, 스토리데이터에 나타난 서사 창작 요소의 분포를 분석하였다. 스토리데이터에

나타난 창작 요소 분포 분석 결과, 스토리데이터 장르에서 ‘Romance’, ‘Drama’가 과반 이상을 차지하며 주제에서도 ‘Love’가 3할 이상을 차지하며 대중문화 콘텐츠에서 사랑과 인간 관계가 중요한 요소임을 확인했다. 스토리데이터의 모티프, 감정 형용사, 행위 동사는 특정 항목에 높은 빈도로 나타났다. 특히 모티프의 경우 빈도가 높은 모티프에서 매우 낮은 모티프까지 분포가 전반적으로 고르면서 매끄러운 이차함수의 모양을 보이면서, 모든 분포에서 상대적으로 빈도가 높은 것이 더욱 높은 향을 보였다. 스토리데이터의 모티프 분포 분석 결과를 통해 대중문화 콘텐츠의 창작자와 소비자의 문화적 선호도를 확인할 수 있었다. 유닛 인과성과 서사단계 분포를 통해서도 대중문화 콘텐츠의 스토리가 주인공을 중심으로 전개되고 있으며, 사건의 추동 및 변동을 이끄는 캐릭터의 주제성을 기반으로 하여 갈등을 해소하고 주인공이 성장하는 과정을 중요하게 나타나고 있음을 확인할 수 있다.

본 논문은 기존의 스토리 데이터베이스 구축이 영문학, 독문학, 한국문학의 고전 소설 및 설화 분석에 집중되어 있으나, 현대의 대중문화 콘텐츠의 스토리를 분석하고 서사 창작 요소를 재창작하여 데이터화 하였다는 의의가 있다. 또한 이러한 대중문화 콘텐츠 스토리 데이터베이스에 나타난 주요 창작 요소의 분포를 분석함으로써 대중문화 콘텐츠에 대한 정량적인 정보를 제공하고자 하였다는 의의가 있으며 본 연구를 통해 구축된 스토리데이터를 이용하여 대중문화 스토리를 추천 및 창작하는 인공지능 모델 개발에 대한 후속 연구가 필요하다.

감사의 글

본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2023년도 문화기술 연구개발 사업으로 수행되었음(과제명 : 인공지능 기반 방송콘텐츠 스토리 분석 및 과급효과 예측 기술개발, 과제번호 : RS-2023-0022907, 기여율: 100%)

참고문헌

- [1] M.-J. Kim, “An Analysis of Search Log from a Story Database Service and a New Story Search Method Based on Story Map,” *Journal of Digital Contents Society*, Vol. 16, No. 5, pp. 795-803, October 2015. <http://doi.org/10.9728/dcs.2015.16.5.795>
- [2] M.-J. Kim, “An Analysis and Visualization of Creative Tendency Appeared in Query Log of a Story Database Service,” *Journal of Korea Multimedia Society*, Vol. 19, No. 8, pp. 1609-1618, August 2016. <http://doi.org/10.9717/kmms.2016.19.8.1609>
- [3] Ewha Womans University, Development of Authoring Tool for Supporting Movie and Animation, Korea Creative Content Agency, Seoul, Project Report (1st year), 2011.
- [4] C.-W. Lee and C.-Y. Lee, “A Study of the Movie & Animation Storytelling Planning and Creative Support System,” *Humanities Contents*, No. 19, pp. 533-557, November 2010.
- [5] C.-G. Lyou and Y. Jeong, “A Study about <The Principle of Causality> of Digital Narrative Tool,” *Humanities Contents*, No. 22, pp. 183-207, September 2011.
- [6] C.-G. Lyou and H.-Y. Yun, “A Comparative Study on the CBR Model of Story Creation Program: Focusing on the <Minstrel> and the <Storyhelper>,” *Journal of Digital Contents Society*, Vol. 13, No. 2, pp. 213-224, June 2012. <https://doi.org/10.9728/dcs.2012.13.2.213>
- [7] M.-R. Choi, Design Strategies for a StoryBank Using a Digital Storytelling Model, Master’s Thesis, Korea University, Seoul, February 2007.
- [8] D.-J. Jeon, Using Digital Storytelling Study of Folk Tale Design Storybank, Master’s Thesis, Andong University, Andong, August 2012.
- [9] E.-J. Kim, D.-I. Shin, and S.-K. Hwang, “A Study on the Development of a Story Database Based on English Literature: Focus on Motif Extracting,” *Journal of Digital Convergence*, Vol. 13, No. 9, pp. 463-472, September 2015. <https://doi.org/10.14400/JDC.2015.13.9.463>
- [10] J.-Y. Lee, A Narrative Theory-Based Story Database of Folk Tales, Ph.D. Dissertation, The Academy of Korean Studies, Seongnam, August 2017.
- [11] S. Lee, “An Idea on Develop Cultural Contents Using the ‘Korean Classics DB’,” *Journal of Korean Classics*, Vol. 59, pp. 129-161, November 2021. <http://doi.org/10.15752/itkc.59.202111.129>
- [12] H.-S. Lee and J.-Y. Kim, “Development of Story Recommendation through Character Web Drama Cliché Analysis,” *The Journal of the Institute of Internet, Broadcasting and Communication*, Vol. 23, No. 4, pp. 17-22, August 2023. <https://doi.org/10.7236/JIIBC.2023.23.4.17>
- [13] National Information Society Agency. NIA [Internet]. Available: <https://www.nia.or.kr/>.
- [14] J.-S. Lee, *Modern Novels and Thematics of Narrative*, Literature and Intellect Series, 2007.
- [15] S. J. Lee, S. H. Yoon, S. B. Lee, Y. H. Kim, W. I. Kim, Y. G. Park, ... and D. R. Jung, “Das Anlegen von “Databases” zu Motiven der deutschen Literatur,” *Bertolt Brecht und das moderne Theater*, No.20, pp. 263-280, February 2009.
- [16] I.-J. Park and K.-H. Min, “Making a List of Korean Emotion Terms and Exploring Dimensions Underlying Them,” *Korean Journal of Social and Personality*

Psychology, Vol. 19, No. 1, pp. 109-129, February 2005.

- [17] S. Sohn, M.-S. Park, J.-E. Park, and J.-H. Sohn, "Korean Emotion Vocabulary: Extraction and Categorization of Feeling Words," *Science of Emotion & Sensibility*, Vol. 15, No. 1, pp. 105-120, March 2012.



김지선 (Ji-Seon Kim)

2021년 : 이화여자대학교
융합콘텐츠학과 (학사)

2022년~현재 : 이화여자대학교 융합콘텐츠학과 석박사
통합과정

※ 관심분야 : 디지털 스토리텔링(Digital Storytelling), 스토리
데이터(Story Data), 인공지능(Artificial
Intelligence) 등



김명준 (Myoung-Jun Kim)

1991년 : 한국과학기술원 (전산학 석사)
1996년 : 한국과학기술원 (전산학 박사)

1996년~1997년: University of Washington Visiting Scholar

1997년~2000년: 한국전자통신연구원 선임연구원

2001년~현재 : 이화여자대학교 융합콘텐츠학과 교수

※ 관심분야 : 컴퓨터그래픽스, 영상처리, 스토리 데이터,
인공지능 등