

# 스토리 검색 서비스의 사용자 기록에 나타난 창작 성향 분석 및 가시화

김 명 준<sup>†</sup>

## An Analysis and Visualization of Creative Tendency appeared in Query Log of a Story Database Service

Myoung-Jun Kim<sup>†</sup>

### ABSTRACT

<Story Helper> is a service providing the story synopses that match user's query. This paper analyzes the user log of <Story Helper> which is the answers to the queries to find stories from database, and shows the tendency distribution of user creation. Specially, we analyze a joint distribution of the genres and actions of stories to get better understanding of the tendencies that cannot be found in the analysis of independent distribution. Furthermore, we define a correlation factor between genre and action, and investigate what combinations of the genres and actions are highly, less, and negatively correlated. Finally, we investigate how the tendencies of characters are related to genres and actions, and propose a visualization method to show the tendencies.

**Key words:** Creative Tendency, Joint Distribution, Story Helper, Story Map, User Log

### 1. 서 론

컴퓨터에 의한 스토리의 생성 기술이 연구되고 있지만 서사 구조가 제한적이며 실제 서사 창작에 사용하기에는 부족한 수준이다[11,12]. 반면 작가의 서사 창작을 실질적으로 지원하는 도구로서 스토리 데이터베이스의 중요성이 강조되어 왔다. 안경진[1]은 데이터베이스 구축을 통한 작가 지원 시스템을 제안하였고, 최미란[2]은 스토리뱅크라는 저작 지원 시스템의 개발을 통해 사용자가 플롯-캐릭터-아이템으로 분기될 수 있는 데이터베이스를 이찬욱, 이채영[3]은 장르, 주제, 인물, 배경으로 유형화한 스토

리 데이터베이스 제안한 바 있다. <스토리헬퍼> [4,5,7,10]는 1,019건의 스토리 아이디어이션(시놉시스) 데이터베이스를 구축하였으며 사용자의 검색에 따라 유사도가 높은 스토리를 보여주는 서비스이다.

본 논문은 <스토리헬퍼> 사용자가 스토리 데이터베이스를 검색하기 위하여 질의어에 답변한 서버 기록을 분석하여 스토리를 구상하는 데 있어 장르나 인물 설정의 선호도와 같은 창작 성향의 분포를 살펴보고자 한다. <스토리헬퍼>의 사용자는 국내 작가나 작가 지망생 등 창작에 관심이 있는 사용자가기에 그 창작 성향을 분석해 보는 것은 의미가 있는 일이다.

※ Corresponding Author : Myoung-Jun Kim, Address: 52, Ewhayeodae-gil, Seodaemun-gu, Seoul 03760 Korea, TEL : +82-2-3277-3347, FAX : +82-2-3277-3359, E-mail : mjkim@ewha.ac.kr

Receipt date : Jul. 28, 2016, Approval date : Aug. 11, 2016

<sup>†</sup> Department of Content Convergence, Ewha Womans University

※ This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIP) (R7117-16-0205, Development of Cloud Collaboration System for Creative Writing)

먼저 각 질의어 별로 사용자 기록에 나타난 성향을 분석하여 기존의 통상적인 영화 스토리의 특성과 다른 점이 있는 지를 살펴본다. 또한, 사용자 기록 통계를 이용하여 영화의 장르 성향 분포와 영화의 행위 성향 분포와의 상관 관계를 분석함으로써 장르와 행위의 연관성을 가시적으로 보여주는 연구를 수행한다. 마지막으로 인물의 창작 성향이 장르와 행위에 따라 다르게 나타나는 지를 살피기 위한 방법으로 스토리맵을 이용한 가시화 기법을 제시한다.

## 2. <스토리헬퍼>의 아이디어이션 서비스 동작 방식

<스토리헬퍼>의 아이디어이션 서비스는 다음과 같이 동작한다. 데이터베이스에는 영화의 인물 관점에서 전개되는 스토리가 개별 데이터로 저장되어 있다. 데이터베이스의 각 스토리 데이터에 대해 그 성향을 구분하기 위해 Table 1에 나타난 29개 질의어에 대한 답변이 작성되었으며, 답변 선택 정보가 해당 스토리의 성향 데이터로서 데이터베이스에 함께 저장되어 있다. 사용자가 Table 1의 질의어에 대해 답변을 하면서 자신이 구상하는 스토리의 성향을 설정하면, 데이터베이스 상에서 유사도가 높은 스토리

Table 1. Query/answers for story search and weighting factors

Query		Answer	WF
Genre		western, SF, gangster, war, melo, drama, romantic comedy, comedy, disaster, erotic, action, society, history, musical, dance, heroism, biography, documentary, experiment, adventure, thriller, horror, fantasy	18
Target		Children, Adult, Family	3
Action		pursuit, adventure, trace, rescue, escape, revenge, puzzle, rival, victim, temptation, transformation, transfiguration, maturation, love, forbidden love, sacrifice, discovery, terrible deed, rise, fall, challenge, pure love	30
Character	Age	infant, teenager, young age, middle age, old age	3
	Gender	male, female, neuter gender	3
	Physical	disabled, unhealthy, normal, healthy, overly healthy	3
	Tribe	human, alien, spirit, animal, non-human	6
	Type	Don Quixote type, Hamlet type	3
	Tendency	outward-looking, introvert	3
	Situated Cognition	sense, intuition	3
	Decision	fact, emotion	3
	Behavior pattern	decision, recognition	3
	Time of interest	past, progress, now, future	6
	Concern of behavior	empathy, act, gain, learning	6
	Response	plan, acting, realization, design	6
	Cause of mental problem	memory, impulsive desire, preconscious, subconsciousness, cognition, consciousness	3
	Class	slave, commoner, nobility, royalty	6
	Economic level	low-income, commonalty, mid-income, high-income, rich	6
	Education	none, low, normal, high	6
	Political status	conservative, neutral, liberal, revolution, anarchism	6
	Relationship	lonely, independent, sociable, dependence	3
	Race	white, black, asian, mixed-blood	3
	Job	elementary, middle, high school student, university student, self-employed, employee, teacher, specialized job, entertainer, daily employee, police, housewife, freelancer, civil servant, criminal, inoccupation, etc	15
	Father-personality	none, indifference, family-oriented, patriarchal, violent	6
	Mother-personality	none, indifference, weak, family-oriented, violent	3
	Childhood	deprived, affluence, intelligent, stupid	3
	Status of marriage	single, married, divorce, remarriage, bereavement	6
	Child	none, yes, distantness or death	3
	Sexual orientation	heterosexual, homosexual, bisexual	3

데이터를 검색하여 추천하는 방식으로 동작한다. 이 때 질의어 마다 사용자 답변 선택이 데이터베이스의 스토리의 답변 선택 항목과 일치하는 경우 Table 1의 가중치만큼의 점수를 부여하고, 모든 점수의 합이 최종 유사도로 계산된다[6,8].

질의어의 대한 답변에는 Table 1에 나타난 것 외에 항상 'noselect'가 있다. 이는 사용자가 그 질의 항목에 대해 답변을 선택 하지 않는 것이다. 이에 대해서는 데이터베이스와 관계없이 해당 가중치를 유사도 점수로 부여한다.

### 3. <스토리헬퍼> 사용자 기록에 나타난 창작 성향

#### 3.1 장르

Fig. 1(a)는 <스토리헬퍼> 사용자가 구상한 스토리 아이디어이션의 장르 비율을 보여준다. 가장 선호되는 장르는 판타지(16%)로 나타났다. 드라마, SF, 스릴러, 멜로, 로맨틱 코미디, 코미디, 액션, 어드벤처가 그 뒤를 이으며, 이들 9개의 주요 장르가 전체의 76%를 차지하였다. 이는 실제 영화에서도 9개 장르가 주를 이루는 것과 일치하는 성향을 보인다. 반면, 실제 영화에서는 9개 장르가 차지하는 비율이 90%에 달하고 특히 드라마에 33%가 편중되어 있는 것과 비교하면 상대적으로 장르들이 고르게 나타난 것이라고 볼 수 있다. 본 논문에서 사용한 실제 영화에서의 장르 비율은 <스토리헬퍼>의 데이터를 기

반으로 계산하였다.

비중이 작은 장르에서도 전쟁, 역사, 사회와 같이 통상적으로 인기가 적은 장르가 3%, 3%, 2%로 실제 영화에서 보다 높은 비율을 보였으며, 에로와 무협이 각각 3%, 2%로 무시할 수 없는 비율을 보이는 특이함도 있었다. 이상의 장르 성향을 살펴보았을 때 작가 및 작가 지망생들이 포함된 사용자층이 스토리의 구상에 있어 다양한 장르를 폭넓게 시도한다는 것과 블록버스터 판타지나 SF에 대한 선호도가 있다고 해석할 수 있다.

#### 3.2 행위 - 이야기의 대상

스토리의 행위, 즉 어떤 행위에 대한 이야기인가를 묻는 질의어는 스토리의 유사도를 결정하는 데 있어서 가장 큰 30%의 가중치를 갖는 중요 항목이다. Fig. 1(b)는 사용자 기록에 나타난 행위 별 비율을 보여준다. 모험, 사랑, 성숙, 추구와 같이 보편적인 것이 50% 가깝게 나타나는 등 일반적인 주제에 대한 선호도를 보였다.

#### 3.3 인물의 특성

Fig. 2는 인물의 특성에 대한 질의어의 답변 별 비율을 나타내는 차트들이다. Noselect가 크지 않은 경우에는 이를 제외하고 표시하였다. 반면, noselect의 비율이 그 의미의 해석이 필요할 정도로 크다고 판단되는 경우에 대해서는 이를 포함하였다.

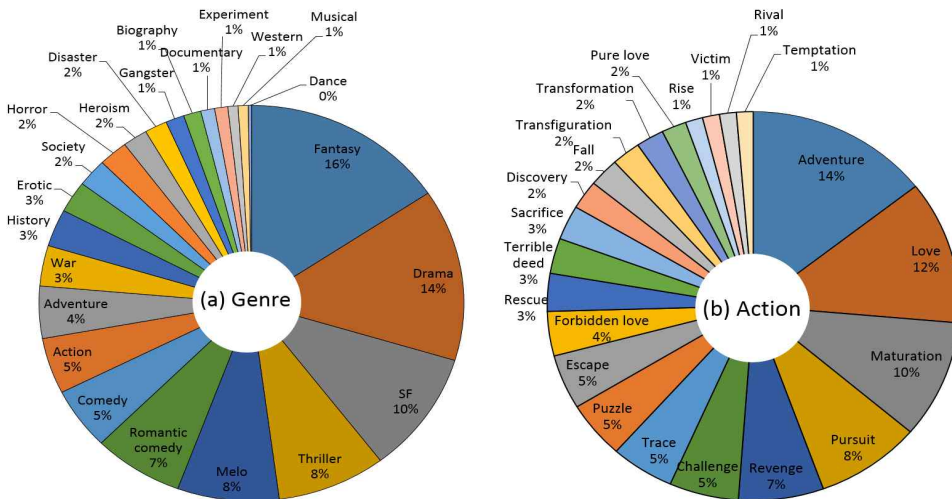
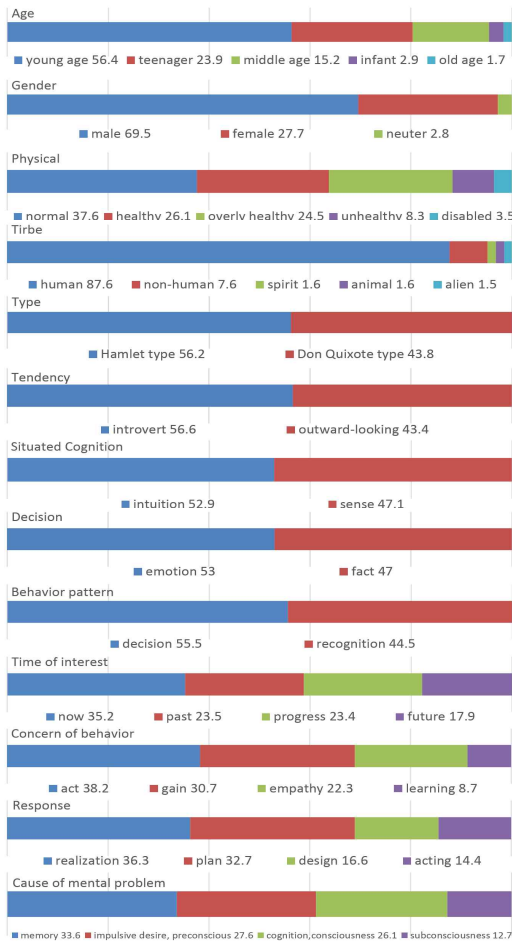


Fig. 1. (a) Genre and (b) Action portions in user log of Story Helper.

### 3.3.1 인물의 연령대, 성별

인물의 연령대 비율은 청년이 56%로 압도적으로 많게 나타났으며 청소년이 그 뒤를 이으며 중장년 보다 많게 나타났다. 실제 영화에서는 청년 및 중장년이 주를 이루는 것에 비해 국내 작가 층의 성향이 좀 더 젊은 인물을 구상하는 경향을 보이는 것으로 해석된다.

인물의 성별은 거의 남자가 여자보다 2배 정도 높은 비율로 나타났는데, 이는 실제 영화에서도 그리 하므로 특이한 사항은 아니라고 볼 수 있다. 특이한 점은 중성의 비율이 3%로 꽤 높다는 것이다. 중성은 로봇, 정성체성 미상, 괴물 등으로 특이한 인물을 설정한 특별한 스토리 구상을 추구하는 작가 층이 상당히 있었음을 의미한다.



### 3.3.2 인물의 신체능력

인물이 신체능력은 보통, 건강, 우월이 주를 이루면서 고르게 나타났고 병약, 장애가 뒤를 이었다. 보통/건강이 평범한 경우로서 이 둘의 비율이 다수를 차지하기는 하였지만, 특별한 경우 즉 우월/병약/장애의 비율이 36%로 실제 영화에서보다 상당히 높은 경향을 보인다. 특히 우월의 비율이 상당히 높다. 이는 <스토리헬퍼> 사용자들이 평범한 인물 보다는 특별한 인물을 설정하는 성향이 높다는 것을 의미하며, 구상 단계의 스토리에서 나타날 수 있는 현상으로 이해된다.

### 3.3.3 인물의 종족

인물의 종족 항목은 인간을 대부분으로 하는 평

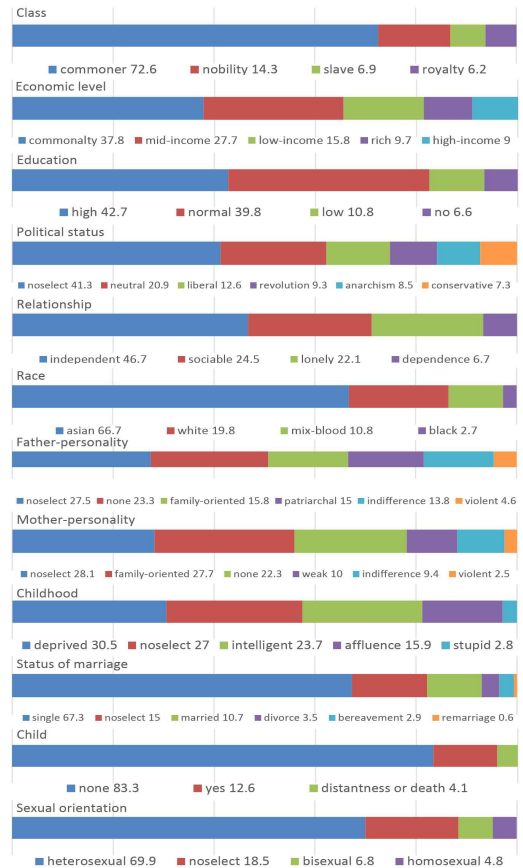


Fig. 2. Creative tendency of characters appeared in user log of Story Helper.

범한 경향을 보였다. 뱀파이어, 늑대인간, 좀비 등을 지칭하는 비인간의 비율이 7.6%로 나타난 점은 재미있는 부분이며 이를 주제로 한 스토리 구상의 비율이 꽤 높음을 의미한다.

### 3.3.4 인물의 유형, 성향, 상황 인지 방법, 판단 방법, 행동 양식

인물의 대한 질의 5~9번은 외향/내향, 직관/감각 등 양자 간의 선택형이다. 이들 질의에 대해서는 거의 50:50의 항목 성향이 나타났다. 즉, 확률적으로 고르게 분포된 것이고 특별한 치우침이 없다고 보인다. 이들 질의어는 매칭률의 중요도 또한 낮은 편으로 가중치가 모두 3%로 지정되어 있다.

### 3.3.5 인물의 관심 시점

인물의 관심 시점은 현재가 약간 더 많지만 과거, 현재, 진행, 미래가 고르게 나타난다. 특별한 해석이 어려운 부분으로 인물 관심사 시점이 구체적으로 설정되지 않아 확률적으로 고르게 나타난 것으로 보인다.

### 3.3.6 인물의 행동 관심사

인물의 행동 관심사는 학습의 비율이 작은 점을 제외하고 고르게 나타났다. 학습 형 인물이 스토리 전개상 선호되는 인물형이 아닌 것으로 인식되었고, 행동 및 생취 같이 능동적인 성향이 주를 이룬다고 볼 수 있다.

### 3.3.7 인물의 문제 대응 방식

인물의 문제 해결 방식 항목에서 특이한 점은 ‘연기’의 비율이 작다는 것이다. 이는 문제 해결을 연기하는 성격의 인물이 좋지 않게 인식되어 낮게 나오는 것으로 보인다. 반면 실제 영화에서는 일은 미루는 성격이라기보다는 우유부단, 고뇌하는 성격 등이 ‘연기’로 지칭되는 것이기에 많이 나올 수 있는 항목이다. 질의어에서 ‘연기’로 지칭되는 인물의 성격의 예를 들어 주었다면, 결과가 다르게 나왔을 것이라 보인다.

### 3.3.8 인물의 계급, 경제 수준, 교육 수준

인물의 계급은 평민인 비율이 70~80%를 차지하였고, 경제 수준은 서민 및 중산층이 각기 30% 정도의 비율을 차지하면서 주를 이루고 저소득층, 고소

득층, 부유층이 나머지를 차지하는 보편적인 경향이 나타났다. 인물의 교육 수준 또한 고등수준과 일반수준의 비율이 80%를 차지하는 보편적인 분포이다. 평범한 계급과 경제, 교육 수준이 다수를 차지하는 자연스러운 분포를 보인 것이다.

### 3.3.9 인물의 정치 성향

인물의 정치 성향 항목은 특별한 경향을 보인다. 사용자 기록에서 noselect가 41%로 매우 많기에 이를 같이 차트에 표현하였다. 이는 인물의 정치적인 성향에 대한 질의에 대해 상당히 많은 사용자가 구체적인 생각을 원치 않았거나 스토리에 불필요한 설정이라 여겼다는 것이다. 또한, 실제 영화에서도 인물의 정치성향이 나타나지 않는 경우가 많기에 인물의 정치적 성향이 매칭률 계산에 영향을 줄 가능성이 적다는 것을 의미한다.

인물의 정치적 성향은 질의어 설계 의도와는 달리 스토리를 선택하는 데 있어 그 중요도가 높지 않다고 할 수 있다. 반면, 이 항목에 대한 가중치는 6%로 인물 항목 중에서는 높은 편이다. 통계적으로 볼 때 매칭의 가능성이 많지 않으므로 가중치를 하향 조정하는 것을 고려해 볼 수 있다.

### 3.3.10 인물의 인종

인물의 인종 비율은 황인이 66%로 압도적으로 높게 나타났다. <스토리헬퍼>의 사용자 층이 국내이기에 자연스러운 현상으로 볼 수 있다. 실제 영화에서는 백인인 경우가 가장 많다.

### 3.3.11 인물의 직업

인물의 직업 분포는 특정 직업이 중요한 의미를 가지는 경우가 적은 것으로 나타나 차트를 표시하지는 않았다. 직업을 지정하지 않은 noselect 및 기타가 36%에 달하고 구체적인 성격의 직업이 아닌 학생이 23%로 나타났다. 연예인, 경찰, 범죄자 등 특별한 의미를 지니는 직업은 2~3%로 소수로 나타나는 특성을 보였다.

### 3.3.12 인물의 부모 성격

인물의 아버지 및 어머니의 성격 항목은 noselect가 많다는 특징을 가지고 있다. 또한 아버지 또는 어머니 ‘없음’이 noselect 다음으로 많다. Noselect

Table 2. Genre–Action joint distribution of creative tendency appeared in user log of Story Helper

(unit: permill %)

	pursuit	adventure	trace	rescue	escape	revenge	puzzle	rival	victim	temptation	transformation	transfiguration	maturation	love	forbidden love	sacrifice	discovery	lembic deed	rise	fall	challenge	pure love
western	1.4	2.1	.8	.1	1.1	1.1	3	3	.0	1	1	1	1	0	0	1	2	1	0	0	2	.1
SF	9.6	22.2	5.5	4.8	7.4	6.4	5.8	.8	1.5	.3	3.0	1.9	6.3	3.6	1.1	4.6	2.9	1.9	1.0	2.0	5.5	.8
gangster	1.6	.6	1.0	.4	.6	2.9	.2	.6	1	.0	2	.4	.7	.7	.2	.0	0	1.1	.5	.9	.8	.2
war	3.2	4.8	1.4	2.8	2.1	3.5	.8	.8	1.0	1	.5	9	1.4	.8	.5	2.3	.3	1.0	.5	1.0	2.4	.4
melo	2.6	1.8	.8	.4	.6	1.7	1.1	.4	.9	1.7	9	9	7.6	34.2	11.7	1.2	.7	1.4	.4	.6	1.1	6.7
drama	12.9	6.2	3.8	2.0	2.9	5.6	4.3	2.5	1.6	1.4	3.2	5.1	33.3	18.9	4.0	4.6	5.5	3.6	2.4	4.8	9.7	2.4
romantic comedy	2.2	3.5	.8	.6	.3	.8	1.3	.5	1	1.6	1.7	1.0	8.0	35.0	3.6	.4	1.1	.6	.7	.1	2.2	3.7
comedy	4.6	8.1	1.9	1.0	1.8	2.0	2.6	1.2	.4	.8	2.0	1.9	5.7	4.4	.6	.8	1.7	.8	1.0	.6	4.6	.6
disaster	1.2	2.4	.6	2.7	3.9	.5	.7	1	.5	.0	5	1	.8	.4	.3	.6	.3	.3	.0	.4	.6	.0
erotic	.4	.9	.2	.0	.0	.4	.2	.6	.2	3.9	.3	.2	1.1	3.9	6.6	1	.0	4.9	.0	.6	.3	.7
action	4.4	7.8	6.4	4.7	1.6	6.9	1.0	.6	.4	.0	1.2	9	2.3	.9	.4	1.7	.6	.8	.2	.8	2.0	.6
society	2.9	1.2	1.8	.4	.9	2.0	1.0	.4	.6	.2	.5	.6	1.9	.4	.4	1.1	.7	.8	.6	1.1	2.7	.1
history	3.9	3.9	1.6	.5	.5	2.3	1.3	.6	.8	1	.6	.4	2.1	1.5	1.6	1.2	1.0	.2	1.3	.8	3.4	.2
musical	.9	.8	.1	.1	.6	.3	1	.3	1	.0	2	.0	1.0	1.3	.2	1	.1	.2	1	.1	.7	.2
dance	.4	.1	.0	.0	.0	.0	.1	.0	.0	.0	.0	.4	1	.0	.0	1	.0	.0	.0	.0	.8	.0
heroism	1.6	4.6	.8	.4	.5	3.3	.5	.5	1	.0	.4	5	1.3	.9	.3	.2	2	.3	.5	.4	1.4	.2
biography	2.5	1.6	.1	.1	.2	.5	.2	1	1	.0	2	5	1.6	.4	.2	.4	.5	.3	.3	.5	2.2	.1
documentary	1.7	.5	.1	.0	.3	1	.2	.0	.3	1	.3	.6	1.7	.5	1	.6	.5	.2	.2	.6	1.4	.0
experiment	.9	.6	.2	.4	.6	.2	.9	1	.4	.3	.3	.2	.7	.3	.2	.2	.6	1.1	.2	.8	.3	.1
adventure	3.7	19.8	1.3	1.4	2.2	1.6	1.8	1	.2	.2	.5	.7	2.4	1.7	.3	.5	1.2	.2	.4	.2	2.5	.4
thriller	3.3	2.2	15.7	3.3	8.6	12.9	12.9	1.0	1.1	1.6	1.3	1.7	2.5	1.4	1.4	1.4	1.1	5.6	.6	3.0	1.0	.3
horror	.6	1.1	1.1	1.4	5.6	1.5	3.2	1	.7	.2	.3	.4	.7	.4	.4	.6	.4	1.6	1	1.2	.3	.1
fantasy	12.6	51.0	4.6	3.9	4.3	10.4	7.3	1.2	2.1	.4	3.7	3.0	13.5	10.0	3.2	5.8	3.4	1.7	2.1	2.7	7.5	1.6

와 ‘없음’의 비율이 50% 넘어 인물의 부모 성격에 대한 설정을 하지 않는 경우가 많았음을 보인다.

사용자 관점에서 ‘없음’은 부모가 없다는 의미인데 부모의 성격을 묻는 질의어에 대한 선택 항목 중의 하나로 조화롭지 않게 보이는 문제가 있다. 또한, noselect는 부모가 있으면서 ‘성격을 지정하지 않음’의 의미인 지 부모가 없을 수도 있는 것인지 혼동의 여지가 있다. 따라서 ‘인물의 부모 있음/없음’을 따로 구별된 질의어로 구성하거나, 인물 부모의 성격을 묻는 항목에서 ‘없음’과 noselect를 하나의 선택으로 통합하는 개선 방안을 고려해볼 수 있다.

### 3.3.13 인물의 어린 시절

인물의 어린 시절은 noselect를 제외하고 보면 ‘불우’가 제일 많고 ‘아둔’이 가장 적은 현상이 나타난다. Noselect가 27%로 매우 많은 비율을 차지하는 특징이 있어, 인물의 정치성향 및 부모 성격과 함께 실질적인 매칭률 기여도가 낮은 항목으로 분류할 수 있다.

### 3.3.14 인물의 혼인 관계, 자녀 유무

인물의 혼인 관계는 미혼이 79%로 압도적으로 많았다. 인물의 자녀 유무는 ‘없음’이 83%로 압도적으로 많은데, 이는 미혼이 많은 것이 거의 그대로 반

영된 것이다. 인물이 미혼이 아니면서 자녀가 없는 비율은 미혼의 비율을 제함으로서 추정할 수 있는데 12%로 계산 된다. 이는 인물의 자녀가 있는 비율과 비슷한 수준으로, 미혼이 아닌 경우 자녀의 유무는 대략 50:50임을 의미한다.

### 3.3.15 인물의 성적 취향

인물의 성적 취향은 이성애자의 비율이 70%로 압도적으로 많고 noselect가 18%로 지정하지 않은 경우도 많았다. 인물의 성적 취향이 양성애자나 동성애자인 비율이 5%가 넘는 상당한 비율을 보였는데, 이는 실제 영화에서의 비율의 크게 상회하는 수준이다. 이는 역시 여러 번 확인 되었듯이 작가 및 예비작가인 사용자들이 좀 더 파격적인 스토리를 구상하는 경우가 많다는 것을 의미한다.

## 4. 장르와 행위의 조인트 분포(joint distribution) 분석

유사도 측정을 위한 질의어 중 영화의 장르와 행위의 가중치는 각각 18, 30으로 그 합이 48로서 매우 높다. 스토리 유사도가 높은 경우가 80% 정도임을 감안하면 장르와 행위가 일치하지 않고는 유사한 스토리로 추천될 수 없다고 할 수 있다. 따라서, 장르와 행위를 스토리의 성향을 결정짓는 구별자로 사용

하여도 될 것이다.

장르의 선택항목은 noselect를 제외하고 23개이고 행위의 선택항목은 22개이다. 따라서  $23 \times 22 = 506$ 개의 장르-행위 조합이 가능하다. 본 절에서는 장르와 행위를 결합한 조인트 분포 (joint distribution)를 살펴보고자 한다. 예를 들어 사용자 기록에서 가장 많은 장르는 판타지로 16%인 데, 이것은 행위에 따라 판타지-추구 1.26%, 판타지-모험 5.1%, 판타지-성숙 1.35% 등으로 장르와 행위가 두 가지가 모두 일치하는 경우의 비율을 계산하는 것이다. Table 2는 사용자 기록에서 나타난 장르-행위의 조인트 분포를 나타낸다. 각 조인트 항목의 비율은 수자가 작은 편이므로 1000분의 1단위인 퍼밀(‰)을 단위로 사용한다.

Table 2의 사용자 기록을 살펴보면 가장 높은 비율은 보인 것은 판타지-모험 51‰이고 로맨틱코미디-사랑 35‰, 멜로-사랑 34.2‰, 드라마-성숙 33.3‰이 그 뒤를 따른다. 장르 분포만을 보았을 때는 로맨틱코미디가 6번째 이지만 장르-행위 조인트 분포에서는 로맨틱코미디-사랑이 2번째이다. 이는 로맨틱코미디 장르 안에서 행위가 사랑으로 쏠려 있어서 나타나는 현상으로 장르, 행위 각각의 분포에서는 볼 수 없었던 성향이다. 행위가 '사랑'일 때, 로맨틱코미디는 35‰로 코미디의 4.4‰의 비해 압도적으로 높다. 또한 행위가 '금지된사랑'일 때도, 로맨틱코

미디는 3.6‰로 코미디의 0.6‰의 비해 역시 압도적으로 높다. 당연하지만 사랑의 행위가 있는 코미디가 로맨틱코미디라는 의미이다.

Table 2의 장르-행위 조인트 분포의 특징은 특정 조합의 항목 비율이 50‰ 넘어서는 반면 과반수의 장르-행위 항목 비율이 1‰에 미치지 못하는 편중 현상이 나타난다. 이는 특별히 잘 어울리거나 서로 맞지 않는 장르-행위 조합이 있어 나타나는 현상이다. 장르-행위 조합의 비율이 낮은 경우는 서로 어울리지 않는 것으로 해석되는 동시에 해당 장르-행위 조합의 스토리가 많이 시도되지 않은 것으로 해석할 수 있어 미개척지로서의 정보의 가치가 있다.

#### 4.1 장르와 행위의 연관성 분석

앞서 서술한 바와 같이 영화의 장르와 행위는 서로 무관한 것이 아니라 특별히 어울리는 또는 선호되는 조합과 그렇지 않는 조합이 있다. 장르와 행위 사이에 연관성이 있는 것이다. Table 2에서 비율이 높은 것이 더 선호도가 높은 장르-행위 조합이라 할 수 있다. 그런데, 이 선호도가 바로 장르-행위 간의 특별한 연관성을 의미하는 것은 아니다. 예를 들어 사용자 기록에서 가장 선호도가 높은 장르인 '판타지'의 비율은 16%이다. 그리고, 가장 높은 행위는 '모험'으로 14%이다. 이 두 가지 장르와 행위는 각자

Table 3. Difference of Genre-Action joint distribution and independent probability distribution, shown only 1‰ or higher (unit: permill ‰)

	pursuit	adventure	trance	rescue	escape	revenge	puzzle	rival	victim	temptation	transformation	transfiguration	maturation	love	forbidden love	sacrifice	discovery	terrible deed	rise	fall	challenge	pure love	
western																							
SF	2	8		2	3		1			-1			-3	-8	-3	2							-1
gangster		-1				2																	
war				2		1							-2	-3		1							
melo	-4	-10	-3	-2	-3	-4	-3							25	9	-1	-1				-1	-3	5
drama	2	-15	-3	-2	-3	-4	-2				2	20	2	-1		2					2	2	2
romantic comedy	-3	-7	-3	-2	-3	-4	-2					1	27	-2		-2	-1				-2	-2	2
comedy						-1								-2	-1								2
disaster				2	3									-2									
erotic	-2	-3	-1			-1	-1			4			-1		6		4					-1	
action			4	3		4	-1						-2	-5	-1								
society	1	-2												-2								1	
history	2													-2								2	
musical																							
dance																							
heroism		2				2								-1									
biography														-1								2	
documentary																							
experiment																							
adventure		13				-1								-2	-4	-1					-1		
thriller	-3	-10	11		5	7	9						-6	-9	-2			3		1	-3	-1	
horror	-1	-2			5		2						-1	-2			1						
fantasy		28	-3	-1	-3					-2				-2	-9	-3	1						-1



독립적으로도 선호도가 높은 것이기 때문에 특별히 연관성이 없다고 가정하여도 판타지-모험의 비율이 높게 나오는 것은 당연할 것이다. 즉, 장르와 행위가 연관성이 없다고 가정한다면 판타지-모험의 조인트 확률은  $16\% \times 14\% = 22.4\%$ 로 계산되며 이것은 높은 확률이다. 다시 말하면 22.4%은 높은 선호도의 확률이지만 이것이 판타지-모험이 특별한 조합임을 말하는 연관성 수치는 아니라는 것이다. 반면 판타지-모험의 조인트 분포 비율이 이보다 높다면 이 조합의 특별한 의미 즉 그만큼의 연관성이 있다고 볼 수 있는 것이다. 실제로 Table 2에서 판타지-모험의 조인트 비율은 51%로 독립적 확률로 계산된 판타지-모험의 조인트 비율 22.4%에 비해 28%이상 높게 나타난다. 즉, 판타지-모험은 매우 연관성이 높은 조합임을 알 수 있다. 같은 방법으로 멜로-모험의 연관성을 계산하여 보면 -10%로 마이너스의 연관성이 높게 나타난다. 이는 멜로 장르와 모험 행위 각각의 선호도의 곱으로 계산된 독립적인 멜로-모험의 선호도 보다 10% 선호도가 떨어지는 어울리지 않는 조합이라는 의미이다.

사용자 기록상의 나타난 장르-행위의 조인트 분포를 이용하여 상기에서 서술한 바와 같이 장르-행위 간의 연관성을 계산하여 Table 3에 나타내었다. 구체적으로는 다음과 같은 수식으로 계산된 것이며 그 절대값이 1% 미만으로 작은 차이를 보이는 것은 나타내지 않았다.

$$(\text{Genre-Action correlation}) = (\text{Genre-Action joint probability}) - (\text{Genre probability}) \times (\text{Action probability}) \quad (1)$$

장르-행위의 연관성이 높을수록 수치가 커지는 것이며 음의 수치는 마이너스의 연관성으로 어울리지 않거나 선호되지 않는 장르-행위 조합을 의미한다. 수치의 절대값이 작거나 없는 것은 특별한 관련성이 없다는 것을 의미한다. 연관성이 높은 장르-행위 조합으로는 판타지-모험 28%, 로맨틱코미디-사랑 27%, 멜로-사랑 26%, 드라마-성숙 20% 등이 나타났다. 반대로 마이너스 연관성이 높게 나타난 조합으로는 드라마-모험 -15%, 멜로-모험 -10%, 스릴러-모험 -10% 등이 나타났다. Table 3의 연관성도표는 스토리 구상 시에 장르-행위의 연관성이 높은 조합을 선택하고자 할 때 참고 자료로 사용될 수 있고, 반대로 연관성이 별로 없거나 마이너스의 연관성을 가지는 조합으로 과격적인 시도를 해 볼 때 참고자료로도 사용할 수도 있다.

### 5. 장르와 행위의 다른 인물 성향 분석

장르와 행위에 대한 창작 성향이 서로 독립적이지 않고 연관성이 있듯이, 인물에 대한 창작 성향 또한 서로 독립적이지 않을 수 있다. 인물의 유형, 성향, 상황 인지 방법, 판단 방법, 행동 양식은 특히 인물의 성격을 지정하는 질의인 데, Fig. 2에서 보는

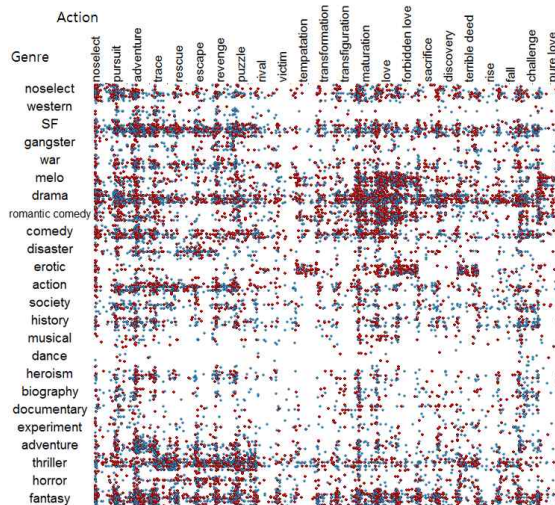


Fig. 3. Creative tendency of character's cognition method shown in a story map. (red: sense, blue: intuition).



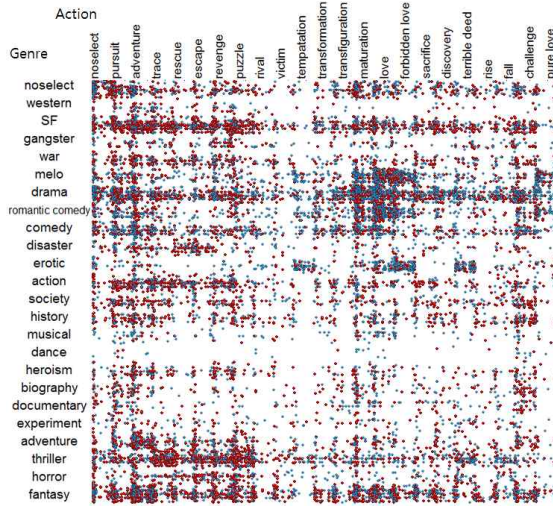


Fig. 4. Creative tendency of character's judgement method shown a story map. (red: fact, blue: emotion).

바와 같이 대부분이 비율이 50:50으로 나타나 개별적으로는 특별한 성향을 보이지 않았다. 본 절에서는 이들 인물의 성격이 장르와 행위에 따라 영향을 받는 지 살펴보고자 한다. 장르 및 행위에 따른 인물의 성향 분포를 가시화하는 방법으로서 스토리맵[8]을 사용한다. 스토리맵은 하나의 스토리를 공간상의 한 점으로 매핑하여 그 분포를 가시화 하는 방법으로서, 행위를 x축, 장르를 y축으로 하여 시작하고 Table 1의 질의어 답변에 따라 계층적으로 좌표를 조정하여 스토리의 성향에 따라 점의 좌표가 달라지게 하는 스토리 성향 분포 가시화 방법이다. 스토리맵이 제안되었던 목적은 데이터베이스의 스토리 분포를 가시화하고 그 중에서 원하는 스토리를 찾아가는 UI이지만, 본 절에는 이를 사용자 기록에 적용하여 창작 성향을 가시화 하는 데 사용한다.

Fig. 3, 4는 사용자 기록에서 나타는 스토리성향을 스토리맵에 나타낸 것이다. Fig. 3은 인물의 상황인지 방법에 따라 감각(적색), 직관(파란색)이 구별되게 하였고, Fig. 4는 인물의 판단 방법에 따라 사실(빨간색), 감정(파란색)으로 구분하였다. 이때, 각 질의어에서 noselect인 부분은 표시하지 않았다. Fig. 2에서처럼 개별적으로 보았을 때에는 별다른 성향을 볼 수 없었던 것에 비해 장르-행위에 따라 인물의 주된 성격이 보이는 부분이 나타났다.

인물의 상황인지 방법은 드라마/멜로-성숙/사랑 부분에서 감각이 우세하게 나타났다. 반면 같은 드

라마 장르라도 드라마-추구/도전에서는 직관이 우세하게 나타남을 보였다. 그 외 직관이 우세하게 나타난 부분은 SF-추구/모험, 어드벤처-모험 등이다. 인물 판단 방법에서는 드라마/멜로 장르 전반에서 감정이 우세한 성향으로 보이면서 드라마/멜로-성숙에서 특히 감정이 크게 두드러진다. 반면 SF는 전반적으로 판단으로 나타났고, 스릴러는 주된 행위인 모험/추적/구출/탈출에서 판단이 강세로 나타났다.

### 6. 결 론

본 논문에서는 <스토리헬퍼>의 사용자 기록에 나타난 사용자들의 창작 성향을 통계적으로 분석하였다. 각각의 질의어 항목에서 나타난 창작 성향은 실제 영화 스토리의 성향과 대체로 유사하지만 실제 영화에 비해 상대적으로 다양한 장르와 특이한 인물 구상을 하는 경향을 보였으며, 블록버스터 판타지나 SF에 대한 선호도가 나타났다. 이는 작가 및 작가 지망생들이 포함된 사용자층이 자신이 희망하는 실험적인 스토리를 다양하게 구상해본 것으로 해석된다.

가장 많은 가중치를 차지하는 장르와 행위에 대해서는 장르-행위 조인트 분포 분석을 수행하여 장르와 행위를 개별적으로 분석할 때는 볼 수 없었던 상세한 성향 분포와 장르-행위의 연관성을 가시적으로 보여주는 연구를 수행하였다. 장르-행위의 조합 단위로 사용자 기록상의 창작 성향을 확인할 수

있었고, 장르와 행위의 독립적 확률을 벗어나는 비율을 그 장르-행위의 특별한 연관성으로 정의하고 연관성이 높은 조합, 연관성이 낮은 조합, 그리고 마이너스의 연관성을 가지는 조합을 가시적으로 확인할 수 있는 방법을 제시하였다.

인물에 관한 창작 성향이 장르와 행위에 따라 특징적인 성향을 보일 수 있어 스토리맵을 응용하여 이를 2차원 공간에 가시화 하는 방법을 제안하였다. 이를 이용하여 특정 장르-행위 조합에서 인물에 관한 창작 성향의 치우침이 나타나는 경우를 확인할 수 있었다.

본 논문은 <스토리헬퍼> 운용에서 생성된 사용자 기록을 바탕으로 국내 사용자들의 창작 성향을 분석한 것으로 2년 이상의 기간에 걸쳐 수집된 충분한 양의 데이터를 기초로 작성되었으며 조인트 분포 분석과 같은 심도 있는 통계 분석을 하였다는 의의가 있다. 향후 보다 효과적인 스토리 관련 정보 서비스의 설계에 기초 자료로 활용될 수 있을 것으로 기대한다.

REFERENCE

[ 1 ] K. Ahn, *Analyzing Writer's Storytelling Task and Designing a Support System*, Master's Thesis of KAIST, 2005.

[ 2 ] M. Choi, *Using Digital Storytelling Study of Folk Tale Design Storybank*, Master's Thesis of Korea University, 2007.

[ 3 ] C. Lee and C. Lee, "A Study of the Movie & Animation Storytelling Planning and Creative Support System," *Humanities Contents*, No. 19, pp. 533-557, 2010.

[ 4 ] Ewha Womans University, *Development of Authoring Tool for Supporting Movie and Animation*, Project Report, KOCCA, 2011.

[ 5 ] C. Lyou and Y. Jeong, "A Study about <The Principle of Causality> of Digital Narrative Tool," *Humanities Contents*, No. 22, pp. 183-207, 2011.

[ 6 ] Ewha Womans University, *Scenario Similarity Analysis System*, 10-1238985, Korea, 2013.

[ 7 ] C. Lyou and H. Yun, "A Comparative Study on the CBR Model of Story Creation Program : focusing on the <Minstrel> and the <Story-helper>," *Journal of Digital Contents Society*, Vol. 13, No. 2, pp. 213-224, 2012.

[ 8 ] M. 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. 797-805, 2015.

[ 9 ] Joint Probability Distribution, [http://wikipedia.org/wiki/Joint\\_probability\\_distribution](http://wikipedia.org/wiki/Joint_probability_distribution), (accessed July 25, 2016).

[ 10 ] Story Helper, <http://www.stotyhelper.co.kr>, (accessed July 25, 2016).

[ 11 ] S. Kim, S. Moon, J. Park, J. Chang, and S. Han, "A Study on Interactive Storytelling Script Language for Generating the Stories," *Journal of Korea Multimedia Society*, Vol. 12, No. 2, pp. 313-322, 2009.

[ 12 ] S.R. Turner, *MINSREL: A Computer Model of Creativity and Storytelling*, PhD Thesis of University of California LA, 1993.



김 명 준

1989년 한국과학기술대학 전산학과 학사  
 1991년 한국과학기술원 전산학과 석사  
 1996년 한국과학기술원 전산학과 박사

1996년~1997년 University of Washington Visiting scholar  
 1997년~2000년 한국 전자 통신 연구원 선임연구원  
 2001년~현재 이화여자대학교 융합콘텐츠학과 교수  
 관심분야 : 컴퓨터그래픽스, 영상처리, 클라우드 컴퓨팅, ICT융합