

# CONCOR 텍스트 마이닝을 활용한 중국 인디게임 현황 분석

## Status and Issue Analysis on Indie Game in China through CONCOR Analysis

최린걸<sup>1</sup>, 김정윤<sup>2\*</sup>

Linje Cui<sup>1</sup>, JungYoon Kim<sup>2\*</sup>

### 요 약

본 연구는 2021년부터 2022년까지의 중국 인디게임 관련 뉴스 빅데이터를 분석해 인디게임의 현황과 이슈를 분석했다. 중국인들이 가장 많이 사용하는 구글, 바이두, 웨이보에 실린 온라인 뉴스 데이터를 추출하고 텍스트 마이닝 사전 처리 과정을 거쳐 키워드를 추출 및 정제하였다. 이후 추출된 키워드를 기반으로 CONCOR 분석을 통해 UNINET과 NET DRAW로 시각화하였다. 분석 결과, 중국 인디게임 산업 발전의 외부 및 내부 요인을 발견했으며 더불어 저해 요소도 발견하였다. 또한 중국 인디게임의 방향은 크게 두 가지로 설명되었다. 첫째, 국내 플랫폼을 기반으로 교육효과를 지닌 인디게임 제작이 필요하다. 둘째, 세계화를 지향하며 해외 플랫폼을 기반으로 '중국적'인 인디게임을 제작해 상품화 노력이 필요하다. 현재 중국의 정책을 보면, 인디게임의 미래는 쉽지 않다. 이는 개인이든 기업이든 게임 관련 국가정책 변화에 주의를 기울여야 하기 때문이다. 또한 미성년자가 게임에 중독되지 않도록 보호하면서 학습의 도구로서 게임을 제작하는 문제를 풀어야 하기 때문이다.

핵심어 : 빅 데이터, 텍스트 마이닝, 중국, 인디게임, CONCOR

### Abstract

This study analyzed the current status and issues of indie games by analyzing big data on online news that were released from 2021 to 2022. Indie game news data from Google, Baidu, and Weibo, which are most frequently used by Chinese, were extracted and refined through a pre-processing of textmining techniques. Subsequently, it was visualized through UNINET and NET DRAW for CONCOR analysis based on the extracted keywords. As a result, external and internal factors of the indie game industry in China were found, and inhibitors were also found. The direction of Chinese indie games was explained in two main ways. First, it is necessary to produce indie games related to children's education. Second, they will strive to produce and commercialize 'Chinese' indie games based on overseas platforms aiming for globalization. Looking at China's current policy, the future of indie games is never easy. This is because Chinese developers, whether individuals or companies, should be careful about changes in national policies.

1 Department of Game and Media, Gachon University, Seongnam, Korea [Graduate Student]

e-mail: 1476450766@qq.com

2 Department of Game Media, College of Future Industry, Gachon University, Seongnam, Korea [Professor]

e-mail: kjyoon@gachon.ac.kr (Corresponding author)

\* 본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 연구개발지원사업으로 수행되었음(과제번호: R2020040243).

\* 본 논문은 2022년도 차세대컨텐츠정보서비스학회 동계학술대회에서 발표한 논문을 수정 및 보완한 것입니다.

Received(December 27, 2022), Review Result(1st: January 15, 2023), Accepted(March 17, 2023), Published(March 31, 2023)



© 2023 The Authors. Published by NCISS.  
This is an open access article licensed under the Creative Commons Attribution-NonCommercial 4.0 International License.  
To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

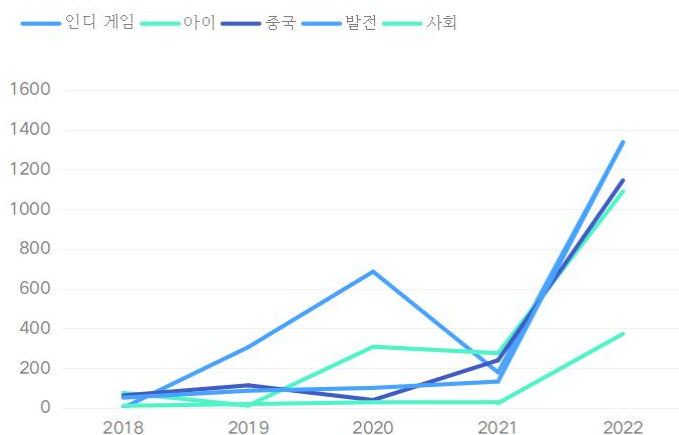
It is also necessary to protect the youth from addiction to games and solve the problem of making games a learning tool.

Keyword : Big Data, Textmining, China, Indi game, CONCOR

## 1. 서론

인디게임(Indie Game/independent Game)은 비디오 게임의 한 갈래로 게임 산업의 절정기에 제기되었으며 영리를 목적으로 개발된 게임과는 다른 개념의 게임이다 [1]. 대형 게임은 자금, 인력, 프로젝트 기간 등 투자를 받아 운영되기 때문에 개발자가 큰 발언권을 가지기 어렵다. 이에 반해 인디게임은 초기 기획 및 개발과정에 있어 외부의 개입이 적기 때문에 개발자가 스스로 게임의 방향을 결정할 수 있고 개발자 본인이 의도하는 바를 표출할 수 있다. 더불어 인디게임이 시장에 진입할 수 있도록 돕는 ‘미니 퍼블리셔’나 스팀을 통해 플레이어 접근성이 높아지면서 인디게임 시장은 점차 성숙기에 들어섰다 [2].

### 키워드 변화

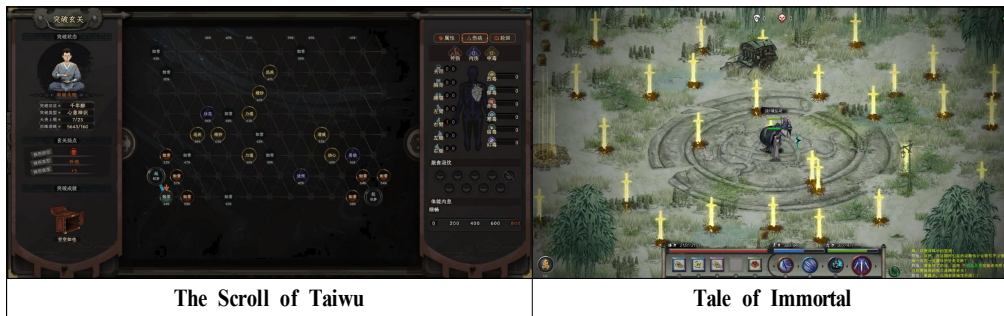


[그림 1] 중국에서 인디게임 관련 온라인 뉴스데이터 양

[Fig. 1] Graph of Online New Data related to Indie Game in China

중국 게임 산업에서 인디게임이 차지하는 비율은 낮다. 하지만 중국의 인디게임은 다양성 측면에서 독창적인 콘텐츠를 제공하고 있기에 중요하다 [3]. 중국 게임업계는 시작부터 정부의 저작권 정책으로 인해 많은 어려움을 겪었다. 그러나 [그림 1]에서 보는 바와 2018년부터 ‘인디게임’ 관련 뉴스가 ‘발전’이라는 어휘와 함께 꾸준히 등장하고 있으며, 외국 플레이어에게도 큰 인기를 얻고 있다. 일례로 [그림 2]와 같이 <The Scroll of Taiwu>과 <Tale of Immortal>은 현지화 번역도 하지 않고 중국어로만 출시되었지만 해외에서 높은 판매량을 거두었다. 정부 정책의 제한에도 불구하고 중국의 인디게임이 발전하고 있는 이유는 사업의 다각화 때문이다. 대기업 산하의 여러 개의 인디

게임 스튜디오가 유동적으로 운영되고 있으며, 필요할 때만 산하 스튜디오를 합작하여 개발하기도 한다. 따라서 중국의 게임 산업의 향후 방향을 이해하기 위해서는 인디게임의 현황을 연구하는 것이 중요하다.



[그림 2] 인기 중국 인디게임

[Fig. 2] Great indie game made in China

중국 인디게임 관련 자료는 포털사이트, 포럼, 잡지, 다큐멘터리 등 각종 매체에 산재되어 있을 뿐 주관적이며 체계적인 정리 또한 부족하다. 또한 대부분 중국 인디게임의 역사를 다루고 있어 산업 현황과 미래 예측자료로 쓰기에는 부족하다 [4][5]. 이와 관련된 보고서도 출간되었으나, 추이를 살펴보는 정도로 국한되어 인디게임의 현황을 살펴보기에 무리가 있다. 이에 본 연구는 연구의 객관성을 높이기 위해 온라인 뉴스데이터를 기반으로 빅데이터 분석 기술을 통해 중국 인디게임 현황을 파악하고 키워드 변화를 통해 이후 방향을 제시하고자 한다. 본 연구는 2021년부터 2022년까지 중국인들이 가장 많이 사용하는 인터넷 검색 사이트인 구글(Google), 바이두(Baidu), 웨이보(Weibo)의 인디게임 관련 온라인 뉴스데이터를 분석할 것이다.

## 2. 이론적 배경

### 2.1 빅데이터 분석

빅데이터 분석이란 복잡하고 큰 데이터 소스를 분석하여 더 나은 비즈니스 결정을 위한 정보를 제공하는 것으로, 트렌드, 패턴, 고객 행동 및 시장 선호도를 발견하는 과정을 말한다. IBM이 제시한 빅데이터의 5가지 특징은 Volume(양), Velocity(속도), Variety(다양성), Value(가치), Veracity(진실성)이다 [6].

### 2.2 텍스트 마이닝(Text Mining)

텍스트 마이닝은 비구조적 텍스트 데이터를 구조화된 자료로 변환하여 정보의 원천으로 활용할

수 있게 하는 기술이다 [7]. 텍스트 발굴을 통해 키워드의 출현 횟수를 추출할 수 있고, 추출한 키워드의 출현 횟수를 분석하여 중요한 데이터의 의미를 파악할 수 있다 [8].

### 2.3 CONCOR(convergence of iterated correlations) 분석

등위성(equivalence)은 네트워크 범위 내의 키워드(노드)들이 서로 유사한 형태의 관계를 맺고 있는지를 나타내는 개념이다. 키워드(노드)들이 서로 동일한 위치에 있을 때 구조적 등위성(Structural Equivalence)이라 한다. 구조적 등위성 집단 내의 키워드(노드)들은 서로 대체 가능하다 [8].

### 2.4 분석도구

자료 수집은 TheIMC에서 제공하는 텍스트(Textom)을 활용하였다. 텍스트는 웹과 SNS에서 대량의 데이터를 실시간으로 수집하고 저장할 수 있으며, 핵심 데이터를 N-gram, TF-IDF, Topic modelling 등의 모델을 통해 빠르고 정확하게 처리하고 정제할 수 있다. 이후 UCI NET, NODE XL 등의 편리한 도구를 통해 다양한 목적에 따라 분석된 내용을 고도로 시각화할 수 있는 도구이다.

## 3. 연구방법

### 3.1 연구 순서

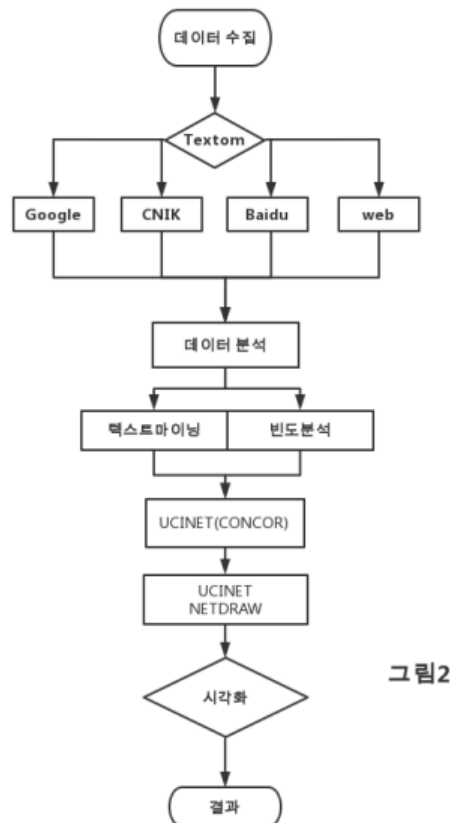
본 연구는 아래의 [그림 3]과 [표 1]에서 보는 바와 같이 텍스트를 통해 관련 자료를 수집하고 텍스트 마이닝을 진행하였으며 핵심 데이터에 대한 빈도 분석을 진행하였다. 또한 UCINET과 NetDraw를 활용하여 결과를 시각화하였다.

[표 1] 연구절차 세부내용

[Table 1] Summary of Online News Data Textmining

단계	내용	
키워드 분석	데이터 수집	온라인 뉴스 데이터 수집 포털사이트: Baidu, Google, Weibo 활용 프로그램: TEXTOM 수집기간 : 2021년 11월 8일~2022년 11월 8일 (총 12개월)
	키워드 추출	키워드 추출 및 정제 응용프로그램: TEXTOM 추출 키워드 개수: 10,481개 방법: 관련 어휘를 통합하여 수정한 후, 100회 이상 인디게임과 관련된 어휘 추출

키워드 분석 시각화	키워드 분석	중요 키워드 분석 활용 프로그램: TEXTOM 방법: 빈도별 상위 핵심 키워드 추출 분석 방법: 빈도분석, TF-IDF 분석, 매트릭스
	의미 연결망 분석	의미 연결망 분석 활용프로그램: UCINET, NETDRAW 분석 방법: CONCOR 분석, 연결중심성 분석



[그림 3] 연구절차 플로우차트

[Fig. 3] Flow chart of Big Data Analysis

### 3.2 데이터 수집

연구에 필요한 데이터 수집은 중국에서 가장 보편적으로 활용되는 검색엔진인 바이두, 구글, 웨이보에서 수집하였으며, 검색어는 ‘인디게임’으로 설정하였다. 또한 인디게임 관련 온라인 뉴스 데이터가 급격하게 증가한 2021과 2022년을 선택하여 중점적으로 분석하였다.

데이터 정제를 위하여 아래의 [표 2]처럼 일부 비관련성 어휘를 배제하고, 같은 종류를 배제하였으며, 같은 의미의 유사 어휘를 병합하고, 최종 추출된 어휘 중 명사만 선택하였다.

[표 2] 데이터 정제

[Table 2] Tokenization of Keywords

데이터 정제	키워드 최적화	키워드 지우기
	‘탐욕동굴(貪婪洞窟)’, ‘기근(飢荒)’, ‘태오회권(太吾繪卷)’ 등 명확한 인디게임명 → 인디게임 ‘동굴(洞窟)’, ‘도서(海島)’, ‘미로(迷宮)’와 같은 인디게임 장르 명칭 → 인디게임 ‘국산(國產)’과 ‘본토(本地產)’, ‘상업(商業)’과 ‘상무(商務)’ → 동의어 정리	‘신작(神作)’, ‘신급(神級)’ 등 인터넷 수식어 → 삭제 게임 및 산업 발전과 상관없는 단어 → 삭제

### 3.3 데이터 분석

본 연구에서는 단어의 빈도수가 100 이상을 넘고 인디게임과 관련된 키워드를 중점적으로 분석하였다. 추가로 빈도수만으로는 키워드의 중요성을 파악할 수 없으므로 키워드의 중요성을 파악할 수 있는 TF-IDF 값과 키워드 간의 연결 중심성 값을 제시하였다. TF-IDF는 여러 문서로 구성된 문서 그룹에서 특정 문서 내에서 어떤 키워드가 얼마나 중요한지 나타내는 통계학적 수치다 [9]. 연결 중심성은 특정 키워드가 다른 키워드와 얼마만큼 직접적인 관계를 맺고 있는지를 계량화한 수치다. 연결 중심성은 UCINET를 활용하여 산출하였다.

## 4. 연구결과

### 4.1 상위 45개 키워드 분석

[표 3]과 같이 2021년부터 2022년 동안 인디게임을 검색 키워드로 설정하고 주요 키워드를 추출한 후 상위 키워드 45개를 선택하였다. 빈도수 순위는 게임(11798), 미성년(1508), 발전(1337), 게이머(1271), 중국(1145), 인디게임(1128), 기업(811), 사용자(787), 플랫폼(754), 혁신성(742), 세계화(560), 업계(547), 개발자(508) 등이다.

TF-IDF의 순위는 미성년(2934.43), 중국(2380.96), 발전(2151.82), 인디게임(2021.11), 기업(1578.13), 사용자(1531.43), 혁신성(1443.87), 플레이어(1396.34), 플랫폼(1350.99), 업종(1311.65), 과학기술(1254.60), 세계화(1230.45), 커뮤니티(1083) 등이다. 연결 중심성은 게임(0.267), 미성년(0.062), 플레이어(0.060), 발달(0.041), 사용자(0.039), 중국(0.0389), 글로벌(0.036), 혁신성(0.0335), 이벤트(0.0334), 인디게임(0.0268), 기업(0.0263), 업종(0.0260), 문화(0.0246) 등 순이었다.

[표 3] 중국 인디게임 중요 키워드 45개

[Table 3] 40 important keywords of indi game in China

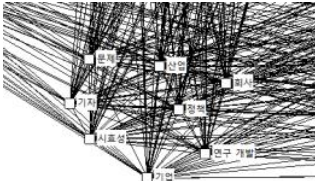
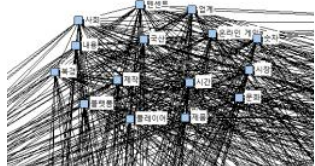
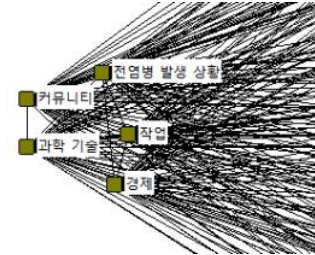
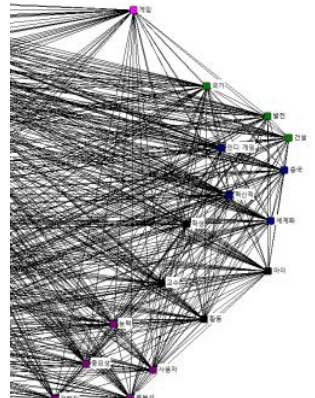
순서	키워드	빈도	TF-IDF	연결중심성
1	게임(遊戲)	11798	0	0.26699
2	미성년(未成年)	1508	2934.43	0.06192
3	발전(發展)	1337	2151.81	0.04118
4	플레이어(玩家)	1271	1396.33	0.06029
5	중국(中國)	1145	2380.96	0.03897
6	인디게임(獨立遊戲)	1128	2021.10	0.02688
7	기업(企業)	811	1578.13	0.02630
8	사용자(使用者)	787	1531.43	0.03907
9	플랫폼(平台)	754	1350.98	0.02198
10	혁신적(創新性)	742	1443.86	0.03350
11	세계화(全球化)	560	1230.44	0.03629
12	업계(企業)	547	1311.64	0.02601
13	개발자(開發者)	508	1056.35	0.01497
14	활동(活動)	482	1059.06	0.03341
15	풍부성(豐富性)	454	997.53	0.01833
16	시장(市場)	434	1145.35	0.02044
17	문화(文化)	412	1087.29	0.02467
18	내용(內容)	393	863.50	0.02371
19	산업(產業)	393	1089.62	0.02630
20	텐센트(騰訊)	383	1061.90	0.01084
21	시간(時間)	382	839.33	0.01651
22	국가(國家)	373	1034.17	0.01843
23	과학기술(科學技術)	362	1254.59	0.01296
24	회사(公司)	348	864.74	0.01987
25	교수(教育)	333	923.27	0.01315
26	정책(政策)	313	886.79	0.01344
27	사회(社會)	301	834.54	0.01680
28	베이징(北京)	293	702.58	0.00988
29	건설(建設)	292	973.00	0.01392
30	온라인게임(網絡遊戲)	289	879.86	0.01920
31	커뮤니티(社區)	283	1206.33	0.01075
32	기자(記着)	280	738.93	0.02361
33	숫자(數字)	278	770.77	0.01132
34	문제(問題)	277	750.12	0.02150
35	능력(能力)	274	723.10	0.01891
36	중요성(重要性)	265	679.71	0.01564
37	제작(制作)	263	694.07	0.00950
38	경쟁(競爭)	259	853.62	0.01056
39	학생(學生)	257	769.90	0.02208
40	국산(國產)	256	709.78	0.00796
41	작업(工作)	253	757.92	0.01939
42	전염병발생상황(疫情)	250	874.12	0.01123
43	제품(作品)	248	716.81	0.01612
44	시효성(時效性)	247	668.88	0.02054
45	연구개발(研發)	231	692.01	0.01084

## 4.2 인디게임 중요 키워드 CONCOR 분석

상위 45개의 키워드의 네트워크 분석 및 시각화를 분석한 결과는 다음의 [표 4]와 같다. 인디게임 키워드를 네트워크 분석하고 유사한 형태의 키워드를 군집화하기 위해 CONCOR 분석을 하였다.

[표 4] 중요 키워드 CONCOR 분석결과 군집별 시각화

[Table 4] Groups of CONCOR analysis with keywords

2021~2022년		
분류	키워드	
인디게임 산업 발전 (8)		문제(問題), 산업(産業), 회사(公司), 기자(記者), 정책(政策), 시호성(時效性), 연구개발(研發), 기업(企業)
인디게임 발전의 내부요소 (15)		텐센트(騰訊), 업계(行業), 사회(社會), 내용(內容), 국산(國產), 온라인 게임(網路遊戲), 숫자(數字), 플랫폼(平台), 플레이어(玩家), 제작(制作), 시간(時間), 제품(作品), 문화(文化), 시장(市場)
인디게임 발전의 외부요소 (5)		커뮤니티(社區), 전염병 발생(疫情), 과학 기술(科技), 작업(工作), 경제(經濟)
인디게임 발전의 장애 요소 (14)		국가(國家), 발전(發展), 건설(建設), 중국(中國), 혁신적(創新), 학생(學生), 세계화(全球化), 미성년(未成年), 교수(敎學), 활동(活動), 능력(能力), 사용자(用戶), 중요성(重要性), 풍부성(豐富性), 개발자(開發者)



CONCOR 분석 결과 군집 1은 기업, 산업, 회사, 정책, 시호성, 연구개발 등으로 형성되어 ‘인디게임 산업의 발전’이라고 명명하였다. 군집 2는 텐센트, 국산, 플레이어, 플랫폼, 시장 등으로 형성되어 ‘인디게임 발전의 내부요소’로 명명하였다. 텐센트는 중국에서 가장 영향력 있는 게임 회사이며 인디게임을 비롯 다양한 장르 게임개발에 주력하고 있다. 또한 국산(중국산) 게임을 제작하고 해외 플랫폼을 통해 유통하기를 바라고 있음을 알 수 있다. 군집 3은 전염병 발생, 경제, 과학기술 등으로 ‘인디게임 발전의 외부요소’로 명명하였다. 코로나19로 인해 중국인들의 격리는 장기화하였고 이로인해 게임을 접할 수 있는 시간이 급격히 증가하였다. 또한 과학기술의 발달도 게임 산업의 발전을 돕고 있다. 군집 4는 학생, 미성년, 중요성, 사용자, 국가, 세계화 등으로 ‘인디게임 발전의 장애 요소’로 명명하였다. 중국 교육제도의 특수성 때문에 중국 부모들은 교육의 실패를 게임 탓으로 돌리는 경우가 많으며, 중국 정책상 미디어 및 콘텐츠 산업에 대한 단속이 엄격하다.

## 5. 결론

본 연구는 2021년부터 2022년까지 인디게임 온라인 뉴스 빅데이터를 분석해 중국 인디게임의 현황과 이슈를 분석했다. 구글, 바이두, 웨이보 웹사이트에 실린 인디게임 뉴스 데이터를 추출하고 텍스트 마이닝 기술을 이용해 키워드 추출 및 정제하였다. 이후 추출된 키워드를 기반으로 CONCOR 분석을 통해 시각화 및 군집화하였다.

분석 결과, 중국의 인디게임 산업이 발전한 외부와 내부 요인과 저해 요소를 발견하였다. 중국 인디게임의 방향은 크게 두 가지로 설명할 수 있다. 첫째, 국내 플랫폼을 기반으로 교육적인 게임 제작이 필요하다. 중국에서는 게임과 미성년자 간의 부정적인 관계를 해결할 새로운 형태의 게임이 필요하며, 둘 사이의 균형을 유지하기 교육적 인디게임이 요구되고 있다. 둘째, 해외 플랫폼을 기반으로 널리 상품화할 ‘중국적’인 인디게임이 필요하다. 과거 중국 게임은 던제 형식이나 고대 중국 신화 및 무협을 다룬 게임이 주종을 이루었다. 중국 게임 관련 고정된 틀을 벗어날 수 있는 돌파구로서 인디게임이 여겨지고 있다.

## References

- [1] M. Gnade, "What Exactly is an Indie Game?", *venturebeat.com*, <https://venturebeat.com/games/what-makes-an-indie-gameindie>, (accessed December 29, 2022).
- [2] K. Graft, "The 5 trends that defined the game industry in 2015", *gamedeveloper.com*, <https://www.gamedeveloper.com/business/the-5-trends-that-defined-the-game-industry-in-2015>, (accessed December 29, 2022).
- [3] D. E. Lee, "The Strategy of Transmedia Storytelling using the Game IP", *Journal of Korean Game Society*, vol. 17, no. 5, October 2017, pp. 143-150, doi: 10.7583/JKGS.2017.17.5.143.
- [4] J. Wilson, "Indie Rocks! Mapping Independent Video Game Design", *Media International Australia*, vol. 115, January 2005, pp. 109-122, doi:10.1177/1329878X0511500111.
- [5] J. V. Bankler, "A chinese view on indie: How is 'indie' defined by professionals within the Chinese indie game industry?", Bachelor Degree, Informatics, University of Skövde, Sweden, 2018. [Online]. [https://www.researchgate.net/publication/360399909\\_A\\_CHINESE\\_VIEW\\_ON\\_INDIE\\_How\\_is\\_'indie'\\_defined\\_by\\_professionals\\_within\\_the\\_Chinese\\_indie\\_game\\_industry\\_Bachelor\\_Degree\\_Project\\_in\\_Informatics\\_30\\_ECTS\\_Spring\\_term\\_2018](https://www.researchgate.net/publication/360399909_A_CHINESE_VIEW_ON_INDIE_How_is_'indie'_defined_by_professionals_within_the_Chinese_indie_game_industry_Bachelor_Degree_Project_in_Informatics_30_ECTS_Spring_term_2018).
- [6] J. S. Perry, "What is big data? More than volum, velocity and variety", *IBM Developer.com*, <https://developer.ibm.com/blogs/what-is-big-data-more-than-volume-velocity-and-variety>, (accessed December 29, 2022).
- [7] S. H. Kim, Y. J. Lee, J. Y. Shin, K. Y. Park, "Text Mining for Economic Analysis", Bank of Korea, Seoul, Korea, WP 2019-18, June 2019. [Online]. Available: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3405781](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3405781).
- [8] J. C. Na, J. Y. Kim, "Analysis of issues in the game industry using big data", *Journal of Digital Art Engineering & Multimedia*, vol. 9, no. 3, September 2022, pp. 299-308, doi: 10.29056/jdaem.2022.09.07.
- [9] J. H. Kim, J. Y. Kim, "Status and Issue Analysis on Indie Game through Big Data Analysis", *Journal of The Korean Society for Computer Game*, vol. 35, no. 2, June 2022, pp. 43-54.