

미스터리 장르 문학의 데이터베이스 설계를 위한 예비적 고찰 - 한국 추리소설을 중심으로*

전세진**

1. 서론
2. 데이터베이스 설계를 위한 유관 자료 검토
3. 기초 데이터베이스 구축 결과 및 통계 분석
 - 3-1. P1 DB: 근대 추리소설의 지형도 탐색
 - 3-2. P2 DB: 추리소설 발행 형식 및 출판 주체의 다양화
4. 관계형 데이터베이스(RDB) 모델링과 장르 라벨링 방법 제안
5. 결론

국문초록

본 연구는 미스터리 장르 문학 중 한국 추리소설을 중심으로 데이터베이스로 구축하고, 그 과정에서 발견된 문제들을 보완하고 장르 문학 연구에 적합한 방식으로 데이터를 재편찬하는 방법을 제안한다. 21세기 이후 한국 미스터리 장르는 비약적인 확장을 하였으나 그 실태를 파악할 수 있는 체계적인 서지 정보는 미비한 상황이다. 이에 본고는 1900년대 초반부

* 이 글은 2025년 11월 15일 대중서사학회 가을 정기 학술대회에서 발표한 논문을 수정·보완하였다.

상세한 심사를 통하여 논의를 발전시킬 수 있게 해주신 심사위원과 논문 작성에 도움을 주신 김바로, 박인성, 유건수 선생님께 깊은 감사의 말씀을 올린다.

** 서울과학기술대학교 인문사회교양학부 초빙교수

터 현재에 이르는 2,200여 종의 서지 정보를 확보해 기초 데이터베이스를 1차 DB로 구축하고, 이를 '한국 추리소설의 시작과 전개기(P1)'와 '미스터리 장르 확장기(P2)'로 나누어 통계 분석을 진행하였다. 분석 결과 P1 시기에는 추리소설이 신문 및 잡지사를 중심으로 유통되었으며, P2 시기에는 자가출판 플랫폼의 등장과 앤솔로지 중심의 출판 방식을 통해 출판 주체가 다양화되는 양상을 확인했다. 이에 이러한 복합적인 서지 정보를 처리하고, 데이터 무결성을 확보할 수 있는 관계형 데이터베이스(RDB) 모델링을 진행하여 2차 DB를 만들고, 하위 장르의 복합적 속성을 반영한 멀티 라벨링 방법을 제안했다.

본 연구는 한국 추리소설의 현황을 객관적으로 확인하고, 장르 문학 연구를 위한 방법론적 기반을 마련하고자 하였다. 제안한 데이터 모델은 향후 텍스트 마이닝 작업과 결합하여, 한국 미스터리 서사 요소적 특징을 규명하는 연구로 확장될 수 있게 되기를 기대한다.

(주제어: 미스터리 장르, 한국 추리소설, 데이터베이스 설계, 디지털 인문학, 관계형 데이터베이스, 장르 라벨링)

1. 서론

21세기 이후 한국의 미스터리 장르는 '탐정·정탐소설' 혹은 '추리소설'이라는 전통적인 문학적 틀을 넘어 비약적인 성장을 거듭해 왔다. SF, 판타지, 스릴러 등 인접 장르와의 활발한 결합을 통해 주제적·형식적 다양성을 확보했을 뿐 아니라, 출판 문학의 경계를 넘어 웹툰, 웹소설, 영상 매체로 확산되며 거대한 장르 생태계를 형성하였다. 특히 웹소설 플랫폼을 통해 등장한 신진 작가들은 기존의 장르 문법을 자유롭게 변주하며 방대한

양의 텍스트를 생산하고 있다.

그러나 “한국 추리소설에는 역사가 없다”는 박진영의 단언처럼, 한국 문학사에서 ‘추리소설’이라는 장르는 오랫동안 문단 안팎에서 비주류 취급을 받으며 그 역사와 가치를 제대로 따져본 적이 없었다.¹⁾ 안혜연이 지적했듯 추리소설은 “고급문화와 대중문화, 순수문학과 대중문학의 이분법 속에서 평가되어 온” 과정에서 김내성, 김성종 등의 대표 작가나 ‘추리소설’ 자체의 개념에 대한 언급조차 적어 한국문학사 내에서 인정받지 못했다.²⁾ 이범봉 역시 “순수문학/비문학의 극단적 분할 구도 속에서 탐정소설은 비문학의 전형으로 간주”되어 “대중적 읽을거리로 소비”되는 경향이 짙었다고 평가한 바 있다.³⁾ 윤영천은 현재 미스터리 카테고리에 포함된 작품들은 대부분 장르의 본질적 특성에서 이탈했기 때문에 ‘범죄 소설’ 정도의 명칭이 적당할 것이라 제시하며, 시대와 국적, 상업적 요구에 따라 끊임없이 변화해 온 미스터리 장르는 모호한 경계를 지니게 되었다고 밝힌다. 그리고 이것이 강력한 대중성 확보와 동시에 명확한 정의를 어렵게 만드는 요인이 되었다고 보았다.⁴⁾ 이러한 논의를 통해 미스터리 장르 문학은 꾸준한 대중적 수요를 통해 발전해왔지만, 미스터리 장르의 개념·작가·작품이 무엇인지를 실증적으로 다루는 연구는 충분히 이루어지지 않은 듯 보인다.

기존의 한국 미스터리 장르에 대한 연구는 세밀한 읽기(close reading)에 집중되어 왔으며, 그 접근 방식은 크게 두 가지 유형으로 구분된다. 첫

1) 박진영, 『탐정의 탄생』, 소명출판, 2018, 6-9쪽.

2) 안혜연, 「1970년대 후반~1980년대 한국 추리소설 붐과 그 사회적 맥락」, 『인문과학』 제85호, 성균관대학교 인문학연구원, 2022, 195쪽.

3) 이범봉, 「탐정물, 1950년대 사회 문화의 숙어」, 『미스테리아』 제19호, 2018, 엘릭시르, 22-23쪽.

4) 윤영천, 『미스터리 가이드북』, 한스미디어, 2021, 277쪽.

번째는 작가론이나 작품 서사적 특징에 관한 연구이다.⁵⁾ 김내성, 김성종을 비롯하여 방인근, 방정환, 김유정 등 주요 작가들이 주된 분석 대상이 되었다. 이러한 연구들은 개별 작가의 작품 세계나 서사 구조 및 기법의 특징을 밝히는 데 기여했다. 두 번째 유형은 시기별 미스터리 장르 소설의 특징과 변화를 중심으로 문학사적 의미를 탐구하는 연구들이다. 근대 초기의 번역 및 변안 소설에 관한 연구를 포함하여, '탐정소설'과 '추리소설' 등의 용어 분화 과정, 그리고 매체 변화가 장르에 미친 영향을 함께 다루는 연구가 이에 해당한다.⁶⁾ 이들 연구는 미스터리 장르의 역사적 계보를 규

5) 광승숙, 「방인근의 탐정소설 연구」, 『Journal of Korean Culture』 제34호, 고려대학교 한국언어문화학술학산연구소, 2016, 69-98쪽; 김영성, 「한국 현대소설의 추리소설적 서사구조 연구」, 한양대학교 박사논문, 2003; 연남경, 「김유정 소설의 추리 서사적 기법 연구」, 『한중인문학연구』 제34호, 한중인문학회, 2011, 55-79쪽; 정종현, 「해방전후 김내성 스파이—탐정 서사의 연속과 비연속—〈태풍〉과 〈청춘극장〉을 중심으로—」, 『현대문학의 연구』 제42호, 한국문학연구학회, 2010, 457-508쪽; 정혜영, 「김내성과 탐정문학」, 『한국현대문학연구』 제20호, 한국현대문학회, 2006; 조은숙, 「탐정소설, 소년과 모험을 떠나다—1920년대 방정환 소년탐정소설의 문학사적 위치와 의의」, 『우리어문연구』 제38호, 우리어문학회, 2010, 615-645쪽; Ekaterina Beglova, 「김성중 추리소설 연구—80년대 장편을 중심으로」, 서울대학교 석사논문, 2020.

6) 김효진, 「한국 근대 추리소설 연구」, 강원대학교 박사학위논문, 2014; 박유희, 「한국 추리서사에 나타난 '탐정' 표상—한국 추리서사의 역사와 이론을 위한 시론—」, 『한민족문화연구』 제31호, 한민족문화학회, 2009, 397-434쪽; 박인성, 『이것은 유해한 장르다』, 나비클럽, 2024; 박진영, 앞의 책; 안혜연, 「1970~80년대 한국 추리소설 연구」, 성균관대학교 박사학위논문, 2022; 오혜진, 「1930년대 한국 추리소설 연구」, 중앙대학교 박사학위논문, 2008; 오혜진, 『대중, 비속한 취미 '추리'에 빠지다』, 소명출판, 2013; 오혜진, 「1950~90년대까지 추리소설의 전개 양상」, 『어문논집』 제44호, 민족어문학회, 2010, 299-320쪽; 이주라, 「근대 초기 번역·변안 추리소설의 수용양상 연구」, 『어문논집』 제61호, 민족어문학회, 2010, 81-109쪽; 이현직, 「한국 근대 추리소설 연구」, 단국대학교 박사학위논문, 2013; 차선일, 「한국 근대 탐정소설 연구」, 경희대학교 박사학위논문, 2012; 최애순, 「1930년대 探偵의 의미 규명과 探偵小説의 특성 연구」, 『동양학』 제42호, 단국대학교 동양학연구원, 2007, 23-42쪽; 최애순, 「〈학원〉의 해외 추리·과학소설의 수용 및 장르 분화 과정」, 『대중서사연구』 제21권 3호, 대중서사학회, 2015, 275-320쪽; 최애순, 「1960년대 서스펜스 스틸러의 유입과 신진 추리소설가의 등장—허문녕, 천불란, 천세옥의 단행본을 중심으로—」, 『현대소설연구』

명하고 작품 목록화 작업을 통해 서지학적 성과를 나타내었다.⁷⁾

그러나 대다수의 연구들이 작품 및 작가 중심의 질적 연구에 집중되었다. 특히 ‘추리소설’이라는 명칭이 어떻게 혼재되며 사용되어 왔는지, ‘추리소설’로 분류되는 작가·작품이 무엇인지, 시기별·작가별 생산량은 어떠한지 등을 파악하기 어려운 실정이다. 이러한 문제를 해소하기 위해서는 우선적으로 한국 문학사에서 ‘추리소설’로 분류되는 작품부터 실증적으로 확인할 수 있는 기초 데이터베이스가 필요하다. 추리소설 연구자들에 의해 축적된 목록과 현재 ‘추리소설’이라는 명칭으로 유통되는 작품들을 비교·분석할 수 있는 토대가 마련된다면, 기존에 제기된 문제들을 실증적으로 검토하고 장르 연구의 방향을 모색할 수 있을 것이다.

이러한 문제의식에서 출발하여, 경계가 유동적인 미스터리 장르 중에서도 ‘추리소설’ 형식의 작품군을 우선적인 데이터베이스 구축 대상으로 삼는다. ‘추리소설’은 근대 초기부터 현재까지 가장 포괄적이면서도 지속적으로 연구되어 온 대상이자, 범주이며, 한국 장르문학 연구에서 일정한 정체성을 유지해 온 개념이기 때문이다.

구체적으로 본 연구는 작가(Author), 제목(Title), 출판사(Publisher), 출간 및 연재시작년도(Start_Year), 연재 종료년도(End_Year), 출판형식(Publication Type), 출처(Source)를 메타데이터 항목으로 설정하여

제89호, 한국현대소설학회, 2023, 247-285쪽; 진설아, 「90년대 추리소설의 장르적 방향성과 한계 연구」, 『문화와 융합』 제44권 1호, 한국문화융합학회, 2022, 351-366쪽.
7) 연구자 별로 ‘추리소설’을 비롯한 미스터리 장르 문학에 관한 서지 자료를 제공하거나 『미스테리아』나 『계간 미스터리』 등의 계간지에서 추리소설사 특집 및 작품 목록 정리 자료를 선보인 바가 있다. 또한 한국추리작가협회(<http://mystery.or.kr>)에서 아카이브를 운영하고 있지만 작가 13명에 대한 정보 외에 작품목록, 수상작 등은 아직 업로드된 것이 없다. 신작소개 항목 중 한국장편에 590편, 한국단편에 85편, 해외신간에 1163편이 책표지, 책소개 글을 포함하여 간략하게 소개되어 있고, ‘리뷰’는 공란이다. (접속일: 2026.1.6.) 본 연구의 DB 구축에서 이 자료들은 활용하지 않았다.

1차 데이터베이스를 구축하여 통계 분석을 진행하고, 2차로 관계형 데이터베이스(RDB) 모델링을 통해 1차 자료의 문제들이 어떻게 보완될 수 있는지를 함께 제시하고자 한다. 먼저 유관 자료의 탐색을 통해 1차 데이터베이스를 크게 두 시기로 나누어 살펴봄으로써, 한국 추리소설의 통시적 변화 양상을 추적하고 각 시기별 특성에 따라 메타데이터 설계와 가중치 설정을 달리하는 유연한 데이터베이스 설계 방향을 제안하고자 한다. 이러한 시도는 한국 추리소설 기초 연구의 현주소를 진단하고, 향후 장르문학 연구의 지평을 확장하는 방법론적 토대가 될 것으로 기대한다.

2. 데이터베이스 설계를 위한 유관 자료 검토

본 연구는 ‘추리소설’을 중심으로 하는 데이터베이스(이후 ‘DB’) 구축을 위하여 DB 구축에 활용된 4가지 유관 자료를 검토하고 DB 구축 과정을 설명한 후, 시기 구분에 관한 논의를 이어가고자 한다.⁸⁾

자료 1) 단행본 자료 - 박진영, 「부록 1 한국 근대 추리소설 목록」, 『탐정의 탄생』, 소명출판, 2018, 376-391쪽.(Source : T)

박진영은 “1900년대 후반부터 1949년까지 발표된 번역·번안소설·창작소설·아동모험소설을 게재 및 발행일자 순으로 정리했으며 추리소설이라

8) 현재까지 구축된 DB와 관련 자료는 모두 하단의 저장소에서 확인할 수 있다.
<https://github.com/SEJIN-JEON/Korean-Mystery-Novel-DB>
data/01_review는 논문 심사본에서 활용된 DB로, 통합본은 엑셀 형태이며 시기 구분한 파일은 csv 형태로 업로드하였다.
data/02_final은 2차 DB로, RDB(관계형 데이터베이스) 방식으로 1차 DB를 재편하며 샘플 모델링을 진행한 것이다. 이에 관한 내용은 4장에서 상술하겠다.

고 보기 어려운 작품, 탐험소설, 첩보소설, 실화나 비화라 하더라도 이 책에서 중요하게 다루어지거나 참조할 가치가 있는 자료를 일부 포함시켰다.”⁹⁾라고 밝히며 1906년 5월 19일부터 당해 12월 31일까지 총 191회 연재된 <신단공안>에서부터 1949년 4월 『신원』에 발표된 김내성의 <벌처기>까지 추리소설 연재본 및 단행본 274종을 소개하였다. 이 자료에서 제공된 서지 정보를 본 DB의 메타데이터 항목에 따라 정리하였다.

이 DB에서 출판형식은 ‘Article’과 ‘Book’으로 나누어진다. 이 자료에는 선집(Anthology)이 없는 반면 신문 혹은 잡지에 단편이나 연재소설로 실린 경우가 다수이고, 그 외에는 단행본 출간이기 때문이다. 단행본 출간 일이나 연재가 시작된 날은 (Start_Year)로 표기되며, 출처에 적힌 대로 연월일을 작성하였고, 연재 종료일(End_Year) 역시 출처 표기에 따랐다. 단행본의 경우 해당란은 공란이 된다. 58번 항목의 ‘구라모치 다카오’는 일본 저자이지만 이 자료에서는 번역 및 번안 소설도 포함이 되어 있기 때문에 동일하게 작성했으며, 84번 항목에 ‘김광배, 단정’으로 되어 있는데 이는 <최후의 승리>의 최초 번안을 ‘김광배(김낭운)’가 시작했다가 김광배의 투병으로 연재 도중 ‘단정’으로 바뀌었기 때문에,¹⁰⁾ 둘을 모두 작성해 둔 것으로 추측하였다. ‘단정’을 김단정이나 단정학 중 하나로 확정할 수 없기에 개별 작가로 구분해둔다면, 이 자료에 실린 작가는 총 121명이 된다.

9) 박진영, 「부록 1 한국 근대 추리소설 목록」, 『탐정의 탄생』, 소명출판, 2018, 375쪽.

상기 자료에서 보면 <신단공안>이 연재된 이후 2007년 창작과 비평사에서 <역주 신단공안>과 같이 단행본으로 발행된 경우가 정리되어 있으나, 본 연구에서는 하단에 단행본으로 발간되거나 선집에 포함되는 경우는 서지사항에 포함시키지 않았다. 연재가 끝난 후에 동일한 이름으로 여러 차례 재발행되거나, 제목이 달라지는 경우, 2000년대 이후에 현대문학선집 및 전집에 수록되는 경우나 다른 제목으로 발행되는 경우가 있기 때문에 통계 처리에 혼선이 생길 수 있기 때문이다.

10) 김영애, 「『최후의 승리』 판본 연구」, 『현대문학의 연구』 제54호, 한국문학연구학회, 2014, 256쪽.

자료 2) 잡지 자료 - 「해방 이후 1950년대(1946~1959)까지 잡지 수록/출간 추리소설」, 『미스테리아』 19호, 2018.07., 43-47쪽. (Source: M)

『미스테리아』 문학동네의 출판그룹 중 하나로 장르문학을 담당하는 엘릭시르에서 2015년 6월부터 격월간지로 발행되어 현재까지 총 58호가 발행되었다. 창간 2주년부터 5년 간 추리소설사 특집을 마련하여 1930년대부터 1990년대까지를 약 10년 단위로 추리소설의 변천사를 다루었다.

순번	권호	발행일	표제
1	13	2017.07	“京城”
2	19	2018.07	1945~1959, 전율의 시대
3	25	2019.07	1960년대, 왜남과 하녀와 간첩의 시대
4	31	2020.07	무너진다, 1970년대—김성종, 김형욱, 무등산타잔, 광주대단지, <수사반장>, 대마초
5	36	2021.07	'80 뭔가 보여드리겠습니다
6	42	2022.07	SPECIAL 1990년대, 포스트모던 크레이지 코리아

그중에서 『미스테리아』 19호에는 1946년부터 1959년까지 잡지에 수록된 추리소설 단편 및 연재 목록과 단행본 출간 목록이 정리되어 있다. 시간 순으로 보자면 방인근의 <괴시체>(영인서관, 1946)부터 <국제살인마>(진문출판사, 1962)까지 총 340종의 서지 정보가 담겨 있다.

순번(No.)은 『미스테리아』 19호의 순서에 따라 잡지 수록 자료를 먼저 작성 후 단행본 자료를 작성한 것이며, 연재 시작일(Start_Year)에 년도와 월이 같이 적힌 경우가 Article이고 그 외에는 Book으로 구분했다. 출판형식과 시작일 및 종료일은 자료 1)과 동일하게 적용되었다. DB에서는 발행 년도의 연속성과 자료 형식을 고려하여 자료 1)과 같은 Sheet에서

처리하였다. 이 중 세 작품(ID: 264, 267, 268)은 동일한 서지사항이므로 출처를 T, M으로 작성하여 통합하였다. 자료 1)과 자료 2)에서 동일한 작품은 총 3편으로 위와 같이 처리되어 있다.

자료 1)과 자료 2)는 1900년대 초반부터 1960년까지를 ‘한국 추리소설 형성과 전개기’를 의미하는 ‘P1(Period 1)’으로 처리되어, 3장에서 상세 분석될 것이다.

자료 3) 온라인 자료 - ‘미스터리 장르’에 속하는 작품 교보문고 Excel 다운로드 및 알라딘 스크래핑 (Source: A, K)

국내 대표 온라인 서적 판매 사이트 네 곳—교보문고, 알라딘, 영풍문고, 예스24—중, ‘교보문고’와 ‘알라딘’에서 ‘미스터리 장르’에 해당하는 서지사항을 수집하였다.¹¹⁾ 이 두 사이트로 한정된 이유는 모두 ‘한국(국내)’ 창작 서적과 ‘미스터리 장르’라는 하위 분류가 있었기 때문이다.¹²⁾ ‘예스24’

11) 하단의 웹사이트에서 목록 확보 및 스크래핑을 진행하였다.

온라인 교보문고 ‘국내 미스터리소설’ :

<https://product.kyobobook.co.kr/category/KOR/010104>

온라인 알라딘 ‘국내 추리소설’ :

<https://www.aladin.co.kr/shop/wbrowse.aspx?BrowseTarget=List&SortOrder=2&ViewType=Detail&ViewRowCount=25&PublishMonth=0&PublishDay=84&UsedShop=0&DiscountPercent=20&BranchType=1&Stockstatus=1&CID=51065>

온라인 교보문고에서는 목록을 ‘카테고리 상품 리스트’라는 명칭 하에 Excel 형태로 일괄적인 다운로드가 가능하며, “순번, 상품코드, 판매상품ID, 상품명, 정가, 판매가, 할인율, 적립율, 적립예정포인트, 인물, 출판사, 발행(출시)일자, 분야”의 항목을 포함한다. 알라딘 웹사이트는 일괄적인 목록 다운로드를 하는 방법은 없으므로 파이썬을 이용하여 저자, 제목, 출판사, 출판년도, 별점 항목을 수집하였다.

12) 교보문고에서 ‘미스터리소설’은 국내도서>소설>장르소설>미스터리/스릴러소설로 분류된다. 장르소설의 하위 유형으로는 SF/과학소설, 판타지소설, 공포/호러소설, 미스터리/스릴러소설, 역사소설, 로맨스소설, 무협소설, 게임소설, 웹소설, 드라마/영화소설 등 총 10개 항목이 있다. 알라딘에서는 국내도서>장르소설>추리/미스터리소설로

의 경우는 ‘국내도서’>장르소설>추리/미스터리장르’로 분류되어 있으나, 번역책과 국내 창작 서적 목록이 통합되어 있었고, ‘영풍문고’는 ‘국내도서’>소설/에세이/시’와 같이 더 단순하게 분류가 되어 있었기 때문에 미스터리 장르의 서적 목록이 따로 구분되어 있지 않았다. DB 구축에서 임의성을 배제하기 위해서는 명확한 출처와 수집 방법이 필요하므로, 교보문고와 알라딘을 기준으로 목록을 확보하는 것이 적합하다고 판단하였다. 이에 따라 교보문고에서는 700종, 알라딘에서는 1,312종의 도서 목록을 확보하였다.¹³⁾ 알라딘은 1976년부터, 교보문고는 박연선의 『여름, 어디선가 시체가』(놀출판사, 2002) 한 권을 제외하고는 2016년부터 검색 시점까지의 도서 목록을 확보할 수 있었다.¹⁴⁾

데이터 처리를 위하여 작가명, 작품명, 출판사명을 동일하게 처리하고, 글과 그림 작가가 병기된 경우 그림 작가는 삭제하였고, 번역책도 목록에서 삭제하였다. 교보문고와 알라딘에서 동일하게 출간된 책은 Source에 모두 적어두었고 ‘큰글씨책’이나 ‘[POD]’와 같이 동일 도서가 여러 형식으로 출간된 경우에도 동일한 제목으로 처리하였다. 자료 형식에서 선집의

분류되며, 하위 유형은 추리/미스터리소설, 라이트 노벨, 판타지/환상문학, 과학소설(SF), 호러·공포소설, 무협소설, 액션/스릴러소설, 로맨스소설 등 총 7개 항목으로 구성된다. 이처럼 국내 주요 온라인 서점에서도 ‘미스터리소설’은 ‘스릴러’(교보문고) 또는 ‘추리소설’(알라딘)과 함께 분류되며 혼용되고 있다. 특히 교보문고에는 ‘추리소설’의 독립 항목이 존재하지 않아 ‘미스터리소설’과 ‘추리소설’의 분류 기준을 파악하기 어렵다. 본 연구에서 구축한 DB를 추후 발전시켜, 각 서점의 분류 체계를 추가로 반영하여 동일 작품이 어떤 범주에 배치되는지 분석한다면, 장르 분류의 비일관성을 실증적으로 드러내는 데 기여할 수 있을 것이라 기대한다. 다만 현재의 연구에서는 해당 내용까지를 포함하지 않았으며 이러한 분류 체계의 불일치와 모호성을 해결하기 위한 라벨링 작업에 대한 제안을 4장에서 제시하고자 한다

13) Excel 파일 다운로드 및 스크래핑한 최초 데이터는 “전처리 전 데이터” 파일에서 확인할 수 있다.(상기 각주 8번 참고.)

14) 본 연구의 기초가 되는 DB 작성을 위하여 데이터를 수집한 2025년 11월 06일까지의 검색 결과를 의미하며, 모든 온라인 자료에 동일하게 적용된다.

형태는 Anthology로, 공동집필을 한 경우는 Book으로 처리하였고, 제목을 일괄적으로 수정하면서 953번 항목과 같이 '큰글씨책'이 교보문고와 알라딘 모두에서 출판되는 경우는 중복되어 A, A, K, K와 같이 최대 4개의 출처가 적히기도 하였다. 이처럼 중복되는 도서를 제외하고 총 1612권의 미스터리 장르 도서 목록을 마련할 수 있었다.

이 자료는 1980년대부터 2025년까지를 대상으로 하는 '미스터리 장르 확장기(P2)'로 구분하여 3장에서 통계 분석을 시도할 것이다.

자료 4) 온라인 자료 - 국립중앙도서관 소장자료 키워드 검색(Source: MK(ModernKoreanNovel), MN(MysteryNovel), DN(DetectiveNovel))

국립중앙도서관 소장자료 고급검색에서 '한국현대소설', '미스터리소설', '추리소설'로 키워드를 조합 검색하여 그 목록을 다운로드하였다.¹⁵⁾ 김성종의 <라인X. 上>을 예로 들어보면, 주제명에 '한국 현대소설'과 '추리소설'이 키워드가 되고, 김다인의 <나비:기억을 지우는 자>(스윙테일, 2011)에서는 주제명에 '한국 현대소설'과 '미스터리 소설'이 키워드로 분류되어 있는 것을 확인할 수 있다.

이를 통해 '추리소설'은 475종, '미스터리소설'은 142종으로 총 617종을 확인하였고, 이중 동일한 목록을 제외하고 총 570종이 확보되었다. 국립중앙도서관 온라인 검색을 통해 확보한 도서 정보는 1982년부터 2025년 검색 시점까지였다. 명확한 출처 표기를 위하여 동일한 메타데이터에 적용하여 DB를 구축하였다.

최대한 많은 서지 목록 확보를 위해 국립중앙도서관 자료를 수집하였으

15) [국립중앙도서관 웹사이트]자료검색>소장자료>고급검색>키워드 '한국현대소설'결과 내 재검색 '추리소설'/'미스터리소설' 이 순서대로 검색하여 목록을 확보하였으며 해당 데이터는 "전처리 전 데이터"에서 확인해볼 수 있다. (각주 8번 참고)

나 이를 3장에서의 통계에 활용하기에 신뢰성이 높다고 판단되지 않았다. 우선 ‘한국현대소설’, ‘추리소설’, ‘미스터리소설’과 같은 키워드로 순차적인 검색을 했을 때, 주제명이 없는 경우나 검색에서 ‘미스터리 소설’이나 ‘추리 소설’이라는 주제명이 있음에도 불구하고 검색에서 누락되는 경우도 다수 존재했다. 또한 서재우의 <눈>은 1, 2권으로 나뉘어져 있음에도 불구하고 “눈: 서재우 미스터리 스릴러 장편소설. 1 - 2(완결)”과 같이, 분권된 경우도 한 편으로 검색되기 때문에 일괄적인 데이터 처리가 어려운 문제가 있다. 또한 이 목록을 P2로 편성하여 활용하려면, 자료 3)의 목록과 중복 검토를 해야 했다. 따라서 이 목록은 이번 연구에서는 정보 공유 차원에서 제공하되 3장의 데이터 통계나 분석할 때에는 포함하지 않았고, 추후 RDB 모델링에서 활용하고자 한다.

지금까지 총 4개의 자료—오프라인 자료 2종—자료 1), 2)—, 온라인 자료 2종—자료 3), 4)—를 검토하였다. 자료 검토 과정에서 확인한 가장 큰 문제점은 근대 이후 현재까지 이어져 온 ‘추리소설’ 장르 전반을 아우르는 통시적 데이터의 부재였다. 특히 온라인 자료에서는 ‘추리소설’, ‘미스터리소설’, ‘스릴러소설’ 등의 용어가 혼재되어 사용되고 있으며, 주요 온라인 서점에서도 이들 용어를 구분하는 명확한 기준을 파악할 수 없었다. 이러한 장르 명칭의 혼용은 데이터의 중복성과 일관성을 확보하기 어려운 요소로 작용한다. 또한 추리소설 목록에 관한 서지 작업이 개별 연구자 및 장르 전문 잡지에 의해 한정된 시기에 선별적으로 진행되고 있어, 근대 초기부터 현대까지를 아우르는 통시적인 서지 정보를 마련하기 어려운 상황이다.

물론 특정 장르만을 분리하여 독립적인 DB를 구축하는 사례는 순문학 범주 내에서도 드문 일이다. 그러나 추후 미스터리라는 거대 장르의 하위 항목들을 정밀하게 분류하고 종합하기 위해서는, 장르적 선명성이 높은 추리소설을 우선적으로 구조화하는 작업이 선행되어야 한다고 판단했다.

이에 본 연구는 정보의 객관성과 데이터의 연속성을 확보하기 위해 출처가 명확한 자료를 중심으로 총 기초 DB를 구성하였다. 구체적으로 자료 1)과 2)를 통합한 'T+M' 시트(P1, 611종)와 자료 3)을 바탕으로 한 시트(P2, 1,612종)를 생성하여, 총 2,223종을 최종 분석 대상으로 확정하였다.

다만 1960년대부터 1980년대 사이의 자료가 부족하여,¹⁶⁾ 본 연구에서는 시기를 두 개로 구분하여 분석을 진행한다. 1900년대 초반부터 1960년까지를 '한국 추리소설 형성과 전개기(P1)'로, 1980년부터 현재까지를 '미스터리 장르 확장 및 다변화기(P2)'로 설정하였다. 이러한 시기 구분은 가용한 데이터의 밀도를 고려함과 동시에, 추리소설의 도입기(P1)와 단행본·앤솔로지 중심의 현대적 변용기(P2)를 대조할 수 있는 구조를 마련한다. 3장에서는 이 구성을 바탕으로 구축된 DB의 메타데이터 항목별 통계적 특징을 분석하고, 장르 경계의 불명확성이나 정보의 산재와 같은 기존 연구의 난제들을 보완할 수 있는 설계 방안을 제시하고자 한다.

16) 1960년대부터 1980년대 사이의 자료가 충분히 확보되지 못한 점은 본 연구의 한계이다. 하지만 이는 출처가 명확하고 일관된 기준을 적용할 수 있는 자료를 우선적으로 선별한 결과이기도 하다. 본 연구는 선행 연구자들이 축적한 기초 자료를 재발굴하고 이를 기계가독형 데이터로 구조화하는 것을 일차적 목적으로 삼았기 때문에, 산발적으로 존재하는 개별 작품 정보보다는 출처를 일괄 적용할 수 있는 목록 자료를 중심으로 DB를 구성하였다. 예를 들어 1974년 한국일보 신춘문예에 당선된 김성종의 <최후의 증인>은 1977년 태종출판사에서 단행본으로 발간되어 국립중앙도서관에 소장되어 있으며 LOD도 부여되어 있지만(KMO000016749), 키워드가 없기 때문에 검색에서 누락되었고, 현재 데이터베이스에 반영되지 못했다. 이처럼 1960-80년대 사이에 대표적 작품들의 존재를 인지하면서도, 목록화가 되지 않아 현 단계에서는 서지 정보를 확보하기 어려운 상황이다. 향후 관계형 데이터베이스 모델링을 통해 DB를 마련한 후, 이 시기의 주요 잡지와 본 데이터베이스에서 확보한 작가들의 작품 연보를 교차 검증하여 공백을 보완하는 방식이 효과적일 것으로 판단된다. 본 연구는 현재 가용 가능한 범위 내에서 데이터베이스 구축의 출발점을 마련하고, 향후 확장을 위한 방법론적 기반을 제시하는 것에 목표를 둔다.

3. 기초 데이터베이스 구축 결과 및 통계 분석

본 장에서는 앞서 구축한 기초 DB를 바탕으로 나눈 두 시기의 메타데이터 항목 별 통계 분석을 진행하고, 이를 통해 한국 추리소설사가 지닌 특성을 고찰한다. 이번 분석은 단순히 수치상의 출판 빈도를 확인하는 수준을 넘어, 특정 작가군으로의 편중 현상, 작가-발행처 간 관계와 특징 등을 데이터의 관점에서 검증해보고자 하는 목적이 있다.

3-1. P1 DB: 근대 추리소설의 지형도 탐색

1) 작가(Author) 분석

P1 DB에서 확인되는 작가는 총 203명이며, 그중 빈도수가 높게 책정된 작가 목록은 다음과 같다.

순서	1	2	3	4	5	6	7	8	9	10
전체 611종 Author	김내성	조능식	방인근	천세욱	N/A	허문녕	현치수	임거문	최유범	조훈파
빈도	54	40	39	39	32	29	23	19	11	9
비율	8.8	6.5	6.4	6.4	5.2	4.7	3.8	3.1	1.8	1.5
Article 504종 기준	조능식	천세욱	김내성	허문녕	현치수	임거문	N/A	방인근	최유범	조훈파
빈도	40	39	35	29	23	18	17	12	11	9
비율	8	7.8	7	5.8	4.6	3.6	3.4	2.4	2.2	1.8
Book 107종 기준	방인근	김내성	N/A	이몽석	이석훈	김교계	송완식	연성흠	노자영	석해조
빈도	27	19	15	5	4	3	2	2	2	2
비율	25.2	17.8	14	4.7	3.7	2.8	1.9	1.9	1.9	1.9

P1 시기의 작가(Author)를 살펴보면, 한국 추리소설사가 특정 소수 작가에 의존해온 지형과 동시에 기존 연구에서 잘 보이지 않았던 존재들을 드러낸다. 김내성이 가장 높은 비중을 차지하며, 조능식, 방인근, 천세욱이 각각 약 6.5%로 그 뒤를 이었고, 작자 미상(N/A)이 5.2%를 차지하였다. 203명의 작가 중 132명이 1회 발표에 그쳤고, 5회 이상 유의미한 수의 작품을 남긴 이는 김내성, 조능식, 방인근, 천세욱, 허문녕 등 5명 내외에 불과하다. 주목할 점은 이러한 데이터가 기존 한국 소설사에서 '추리소설'이 겪어온 주변화를 확인하게 해준다는 점이다. 그간의 한국 근대문학사나 소설사에서 '추리소설'이나 '김내성'이 아예 언급되지 않는 경우가 많았다.¹⁷⁾ 또한 개별 연구자에 의해 주로 '김내성' 개인 혹은 '방인근', '천세욱' 등 일부 작가에 국한되어 왔지만 본 DB에서는 2회 이상 작품을 발표한 71명의 작가를 비롯해 필명('꿈길', 'SI생', '봄바람' 등) 뒤에 숨은 다수의 창작 주체들을 발견했다. 게다가 상위권에 집계된 '작자 미상'과 가명의 존재는 장르의 추리소설의 사회적 위치를 반영한다. 당대 양주동, 김유정 등 주요 문인들이 추리소설 변안·번역에 참여하면서도 대중 오락물로 치부되던 장르적 시선 때문에 본명을 감추었던 정황이 데이터상 'N/A'나 '필명'(1회 기록)으로 나타난 것이다.¹⁸⁾ 이처럼 데이터 중심의 접근은 그간 문학사에

17) 『한국문학통사 5(근대문학 제1기)』(조동일, 지식산업사, 2005), 『한국현대문학사』(권영민, 민음사, 2002)나 『한국소설사』(김윤식·정호웅, 문학동네, 2000) 등 대표적인 문학 및 소설사 자료에서도 '추리소설'이나 '김내성' 등의 작가에 대해서는 언급된 바 없으며 다른 작가들의 경우에도 특정 동인에 이름을 올린 경우를 제외하면 '추리소설'이라는 장르와 관련한 언급은 거의 찾아볼 수 없다.

18) 한이, 「한국 미스터리 흥행의 어제와 오늘」, 윤영천 편, 『미스터리 가이드북 - 한 권으로 살펴보는 미스터리 장르의 모든 것』, 한스미디어, 2021, 264쪽.

한이에 따르면, 1920년대 탐정소설은 교양 있는 청년층과 학생들을 대상으로 한 지적 작품이라는 점이 강조되었음에도 불구하고, 실제 번역과 창작에 참여했던 문인들은 탐정소설 작가라는 정체성을 드러내길 꺼렸다고 한다. 양주동, 이하운, 김환태, 김광섭, 이현구, 김유정 등 당대 문인들이 추리소설 변안 작업을 했지만, 이 장르가 저급한

서 가시화되지 않았던 주체들을 포착해내는 것이 가능해졌다.

신문 및 잡지 연재나 단편을 의미하는 Article을 중심으로 보면 조능식<천세옥>김내성 순이며, 단행본 발간 기준으로 보면 방인근<김내성>작자 미상 순이다. 김내성이 Article과 Book 모두에 상위권에 위치하는 반면, 조능식은 단행본 발간이 거의 이루어지지 않았으며, 오히려 방인근은 단편이나 연재본을 발표하기보다 단행본 중심의 발간이 이루어졌다는 사실이 확인된다. 이는 조능식이 주로 번역가로서 활동하며 잡지 지면을 통해 번안 작품을 생산했기 때문으로 추정되며, 방인근 역시 소설 번역도 다수 진행했으나 <마도의 향불>(1932, 『동아일보』 연재 후 단행본 발행)나 명탐정 '장비호'를 등장시킨 탐정소설 연작들(1948)이 흥행하면서,¹⁹⁾ 단행본 중심의 발간 체제를 구축했던 것으로 보인다.

2) 발행처(Publication) 분석

이 시기에 '발행처(Publication)'로 정리되는 신문사, 잡지사, 서적 출판사는 총 134종이며, 이중 10회 이상은 발표된 것으로 집계되는 출판사는 17종, 2~8회가 45종, 1회가 72종이다. 전체 611개 항목 중 Article로

대중 오락물로 여겨지던 사회 분위기 속에서 분명 대신 해몽생, 피피생, 불바람 같은 가명을 사용해 신분을 감추려고 하였기 때문이다. 이러한 이름들은 추후 RDB로 DB를 재편하는 과정에서 통합한다면, 실재했던 작가적 정체성을 복원하는 과정에서도 그 기능을 수행할 수 있게 된다.

19) 한이는 한국 미스터리 장르의 역사를 돌아보면서 '방인근'에 대해 "이 시기 눈에 띄는 작가는 방인근인데, 원한의 복수, 국보와 괴적, 괴사체 등의 작품에서 탐정 캐릭터인 '장비호'를 등장시켜 사건을 해결한다. 순수 창작이었다는 점과 한국적 현실에 맞춘 추리소설이었다는 점에서 높이 살만하지만, 추리소설의 위상을 높였다고 보기는 어렵다."고 단평한 바 있다.(한이, 「한국 미스터리 흥행의 어제와 오늘」, 앞의 글, 267쪽.) 1950~60년대에 '허문영'도 장르작가로 일부 연구되었으나 여전히 그 양적·질적 연구의 확장이 필요한 상태라 하겠다.(대중서사장르연구회, 『대중서사장르의 모든 것 3: 추리물』, 이론과 실천, 2011, 84쪽.)

분류된 항목이 504종이며 이에 해당하는 출판사는 72종, Book으로 분류된 항목은 107종이며 이에 해당하는 출판사는 62종이다.²⁰⁾ 각각 순차적으로 통계를 내면 아래와 같다.

순서	1	2	3	4	5	6	7	8	9	10	11	12
전체 134종 출판사	아리랑	별건곤	실화	화제	야담	명랑	소설계	이야기	조광	동아일보	야담과 실화	학원
횟수	88	37	36	36	31	21	18	18	16	13	13	13
비율	14.40	6.06	5.89	5.89	5.07	3.44	2.95	2.95	2.62	2.13	2.13	2.13
Article 504종	아리랑	별건곤	실화	화제	야담	명랑	소설계	이야기	조광	동아일보	야담과 실화	학원
횟수	88	37	36	36	31	21	18	18	16	13	13	13
비율	17.5	7.3	7.1	7.1	6.2	4.2	3.6	3.6	3.2	2.6	2.6	2.6
Book 107종	영장서관	문연당	동양 서원	청운사	박문 서관	대지사	조광사	문생당	청록 문화사			
횟수	6	6	5	5	4	4	3	3	3			
비율	5.6	5.6	4.7	4.7	3.7	3.7	2.8	2.8	2.8			

Article의 경우에는 ‘아리랑’이 가장 높은 비율을 차지하고 별건곤>실화>화제 순으로 집계된다.²¹⁾ 현재 DB에서는 전체 611종 중 504종이

20) 현재 구축된 기초 데이터베이스에서는 연재본의 횟수에 대한 개별 가중치를 설정하지 않고 단행본과 동일한 단위로 처리하였다는 분석상의 제약이 존재한다. 이는 향후 보완 과정에서 각 연재물의 구체적인 회차 정보와 발행처 사이의 관계를 가중치로 반영할 수 있는 데이터베이스를 설계함으로써 해결할 수 있을 것이다. 다만 본 연구의 RDB 모델링에서는 엔솔로지를 중심으로 작가의 가중치 차이만 두었다. 연재 일자와 횟수를 파악하여 이 작업을 진행하는 것은 향후 과제로 두고자 한다.

21) 잡지 『별건곤』(개벽사, 1926-1934, 통권 74호)과 『아리랑』(아리랑사/삼중당, 1955-1980), 『실화』(창간-폐간 호수 확인 X)은 대상과 풍속, 통속적 소재를 다루며 폭넓은 독자층을 확보했던 대중잡지이다. 『야담』(야담사, 1935-1945)과 『야담과 실화』(야담과 실화사, 1957-1958)는 설화·괴담·기담 등 야담 장르를 전면에 내세운 야담 전문 잡지로, 탐정소설과 명랑소설을 비롯해 독자의 호기심과 오락성을 자극하는 연재물이 집중적으로 실린 매체였다. 이러한 잡지들은 사실성과 자극성을 결합한 사

Article에 해당하므로 그 비율이 유사하며, 상대적으로 단행본(Book)의 수가 적기 때문에 온전한 비교는 쉽지 않지만, 영창서관=문운당>동양서원>청운사 순으로 집계되며 ‘아리랑’과 같이 특별히 높은 비율로 출판이 이루어지기보다는 고르게 분포하는 것으로 확인이 된다. 다만 상기 작가(Author)의 경우와 마찬가지로, 이 출판사의 관계도 단순히 이름만으로 분류하는 것에는 무리가 있다. 예를 들면 ‘조광’은 조선일보에서 창간한 잡지이다. 위 표에는 작성되어 있지 않지만 13위가 조선일보로 11회인데, 이런 신문사와 신문사 창간 잡지의 관계를 어떻게 볼 것인지도 따져볼 필요가 있기 때문이다. 동아일보(13회)와 신동아(8회)도 유사하다. 또한 Article의 경우에는 1900년대부터 1960년대까지를 일괄적으로 적용하기보다는 시대별로 매체가 새롭게 생겨나고 사라지는 경우가 있기 때문에 조금 더 시기를 나누어 확인할 필요도 있다.

구분	1900-1910	횟수	1920	횟수	1930	횟수	1940	횟수	1950	횟수
1	매일신보	4	청년	11	별건곤	34	아이생활	5	아리랑	88
2	황성신문	1	동아일보	7	조광	14	농업조선	3	실화	36
3	제국신문	1	동명	5	월간매신	10	조광	2	화제	36
4	태서문에 신보	1	어린이	4	신동아	8	매일신보 조선일보 과학조선 문에 신원	1	야담	31
5	녹성	1	학생계 조선일보	3	조선일보	7			명랑	21
6			매일신보 별건곤 중외일보		동아일보 사해공론	5				

건 기록, 통속소설, 삽화 등을 통해 한국전쟁 전후시기 대중문화 취향을 드러내는 중요한 자료로 꼽힌다.(공임순, 「1950년대 『야담과 실화』 필화 사건과 두 하위 장르의 선정성 논란」, 『대중서사연구』 제31권 2호, 대중서사학회, 2025, 167-204쪽 참고.)

1900년대부터 1950년대까지 10년 단위로 나누어 살펴볼 때, 신생 출판사(신문 및 잡지)의 추이를 확인할 수 있다. 초기에는 '신문사' 위주의 연재물이 중심이었다가, 1920년대부터는 잡지의 비중이 높아지기 시작하다가, 1930년대에는 『별건곤』, 『조광』과 같은 대중잡지의 등장의 역할이 컸음이 확인된다. 1940년대에는 아동잡지 『아이생활』이 상위권에 있지만 전체적으로 발행 횟수가 적었는데, 일제 말기부터 해방기의 혼란한 사회상에서 겨우 명맥을 이어왔다고 추측해볼 수 있다.

순서	출판년도	횟수	순서	출판년도	횟수
1	1957	90	6	1933	20
2	1958	78	7	1935	16
3	1956	64	8	1932	14
4	1959	64	9	1930	13
5	1934	23	10	1949	13

1950년대에는 출판 횟수나 연재물의 양적 증가가 확연히 포착된다. 해방기와 한국전쟁기를 지나 문화 재편기가 도래함과 동시에 장르물 전문 잡지들이 창간된 것도 원인으로 꼽을 수 있다. 박광규는 1950년대를 한국 추리소설의 공백기로 인식하는 경우가 많다. '김내성에서 김성종으로 바로 이어진다'는 식의 단순한 설명이 이러한 오해를 심화시켰다고 보면서, 실제로는 1950년대에도 상당량의 추리소설 작품들이 발표되었음을 밝힌 바 있다.²²⁾ 이봉범도 1950년대 미국 원조를 통해 출판과 미디어 등 문화 기반이 복원되면서, 탐정소설은 명랑소설과 함께 대중문학의 주류로 부상

22) 박광규, 「1950년대 한국 추리소설사의 조각 맞추기」, 『미스터리』 제19호, 엘릭시르, 35쪽.

했음을 지적하며 탐정소설이 문학사적으로 순수문학 진영의 배척에도 불구하고 명맥을 이어올 수 있었던 이유를 분석하였다.²³⁾ 본 DB 분석 결과, 1930년대 초중반과 1950년대 후반의 높은 작품 빈도수는 ‘잡지’ 매체의 활성화와 밀접한 관련이 있으며, 이는 한국 미스터리 장르사가 특정 매체 환경과 긴밀하게 연동되어 전개되었음을 보여준다.

3) 작가-발행처 네트워크 관계 패턴

이 절에서는 P1에서 나타난 작가와 출판사의 관계 패턴을 분석해보고자 한다. 특정 매체 집중형 작가, 다수 매체 협업형 작가, 매체 집중도를 확인해보면 다음과 같다.

① 작가-발행처 페어 빈도수 분석

순위	작가 (Author)	발행처 (Publication)	페어 빈도수	전체 작품 수 대비 비율
1	천세욱	아리랑	15	38.5%(15/39)
2	조능식	아리랑	13	32.5%(13/40)
3	임거문	아리랑	12	66.7%(12/18)
4	최유범	별건곤	10	90.9%(10/11)
5	현치수	야담	10	43.5%(10/23)
6	조훈파	야담	9	100%(9/9)

Article을 기준으로 작가별로 어떤 발행처에 작품을 실었는지를 확인한 결과이다. 한 작가가 특정 발행처에 게재된 순서대로 정렬하여 5순위까지를 정렬하면 상기 표와 같다. ‘아리랑’이 가장 많은 발행을 한 곳이기 때문

23) 이봉범, 「탐정물, 1950년대 사회 문화의 속어」, 『미스터리』 제19호, 엘릭시르, 22-23쪽.(목차에는 「1950년대 대중잡지의 현황」으로 표기됨.)

에 상대적으로 ‘아리랑’에 많은 작가가 연결되는 것은 당연해보이나, 당대 작가 중 천세옥은 ‘학원’(6회), ‘화제’(5회), ‘이야기, 실화’(3회), ‘야담과 실화’, ‘소설계’(2회), ‘명랑’, ‘비밀’, ‘혜성’(1회)인 것을 보면 ‘아리랑’에 특별히 많은 작품을 게재한 것을 추적하여 확인할 수 있다. 그렇다면 천세옥의 작품 경향은 ‘아리랑’과 확실한 협업이 있었다는 것을 입증할 수 있다. 최유범의 경우는 이 시기에 ‘별건곤’(10회) 외에 ‘아이생활’(1회)에만 발표를 하였으므로 주로 ‘별건곤’에만 작품을 실었음이 확인된다. 조훈파의 경우는 전부 ‘야담’에만 발표했으므로 특별한 관계로 이해해볼 수 있다.

② 다수 발행처 협업형 작가

	작가(Author)	연결 발행처	전체작품 수	최대 기고 발행처
1	김내성	33	54	농업조선, 청운사(5회)
2	방인근	24	39	문운당(6회)
3	허문녕	11	29	소설계(8회)
4	천세옥	10	39	아리랑(15회)
5	조능식	6	40	아리랑(8회)

이 그래프는 Article과 Book 전체 통계를 활용한 것으로, 김내성이 총 33개의 발행처에 작품을 실은 것으로 집계되었다. 단행본 출판사(청운사 문성당 등)와 잡지(농업조선, 신여성 등)를 망라하여 김내성의 작품이 첫 발간된 이후부터 꾸준히 미스터리 시장 전반에 걸쳐 유통되었고 폭넓은 활동을 했다는 것이 확인된다. 그 다음으로 ‘방인근’은 총 24개 발행처에 기고가 되었는데, ‘문운당’(6회), ‘아리랑’·‘야담’·‘대지사’(3회), ‘동아일보’, ‘문성당’·‘범조사’·‘대문사’(2회), 그 외 16개의 발행처에는 1회씩만 발표 및 출간한 것이 확인된다. 방인근은 주로 단행본의 형태로 출간을 많이 했

으며 보다 다양한 발행처에 작품을 실었다는 점에서 출판계 전반에서 활동을 했다는 점이 인상적이다. 특정 매체에 의존하지 않고 지속적인 출판을 공략한 것으로 해석된다. 허문녕과 조능식은 주로 잡지에만 기고를 했으며, 천세옥은 잡지 및 신문에 고르게 발표하였다. 앞서 ①번 항목에서 ‘조흔파’가 ‘야담’에만 발표한 것과 달리, 여러 발행처에 기고한 작가들의 패턴을 종합하면, 1950년대 미스터리 장르 시장은 김내성과 같은 저명 작가가 다양한 매체를 통해 폭넓은 독자층을 확보하는 한편, 천세옥, 최유범 등 특정 잡지(아리랑, 별건곤 등)에 집중적으로 기고하는 작가군이 공존하는 구조였음을 확인할 수 있다.

이로써 1900년대 초반부터 1960년까지의 기초 데이터베이스(P1)를 바탕으로 근대 추리소설의 작가 및 발행처 지형을 분석해보았다. P1 시기의 서지 데이터는 당시 추리소설이 특정 매체와 밀접하게 결합하여 생산·유통되었는지를 보여준다. 또한 작가 분석 결과, 김내성과 같은 소수의 대표 작가 외에도 기존 문학사에서 충분히 조명되지 않았던 조능식, 천세옥, 허문녕 등 다수의 작가가 유의미한 활동을 전개했음이 확인되었다. 특히 이들의 활동 기반이 되었던 발행처를 분석해 보면, 당시 대중의 일상적 오락을 책임졌던 통속 잡지가 높은 비중을 차지한다. 이는 근대 추리소설이 잡지 연재를 거쳐 단행본으로 이어지는 선순환 구조를 가졌거나, 혹은 잡지라는 매체 환경에 고착되어 있었음을 의미한다. 또한 작가와 발행처 간의 관계를 살펴본 결과, 특정 매체에 전속되다시피 활동한 작가군과 여러 발행처에 발표했던 작가군이 대비되었다. 이러한 데이터 기반의 지형도 파악은 단순한 수치 제시를 넘어, 그간 질적 연구만으로 파악하기 어려웠던 근대 추리소설 생태계의 실제적 연결망을 복원할 수 있음을 확인하였다. 이어 3장 2절에서는 P2 DB 분석을 이어간다.

P2 DB를 분석하기에 앞서 이 두 DB 사이에 차이가 생기는 지점이 출판

형식(Publication_Type)인 점을 강조하고자 한다. P1의 경우 개별 작가에 의해 생산된 작품이 신문이나 잡지의 연재본으로 발표되거나 단행본으로 출간된 사례가 대부분을 차지하는 반면, P2에서는 단행본 출간을 기본 단위로 하되, 여러 작가의 작품이 함께 수록된 앤솔로지 형태의 저서가 상당한 비중을 차지한다. 이러한 형식적 차이는 작품 수, 작가 수, 출판 빈도와 같은 기초 통계 지표에 영향을 미치며, 동일한 기준으로 두 시기의 데이터를 비교할 경우 특정 시기의 특성이 과대 혹은 과소 대표될 위험을 내포한다. 이러한 문제는 4장에서 보완 방향을 언급할 것이며, 이를 유의하며 P2 DB를 살펴보도록 하겠다.

3-2. P2 DB: 추리소설 발행 형식 및 출판 주체의 다양화

1) 작가(Author) 분석

P2 DB는 출판 형식을 단행본(Book)과 앤솔로지(Anthology)로 나누어 정리하였으며, 총 1,612종의 데이터 중 단행본이 1,477종, 앤솔러지가 135종을 차지한다. P2 시기의 작가 지형은 김성종으로 대변되는 ‘스타 작가의 탄생’과 앤솔러지를 매개로 한 ‘작가군 네트워크의 형성’이라는 두 가지 축으로 요약된다.

순위	작가 전체	횟수	엔솔로지	횟수	단행본	횟수
1	김성중	108	정명섭	25	김성중	102
2	정명섭	50	조동신	23	정명섭	25
3	정해연	40	황세연	20	정해연	23
4	김재희	39	한국추리작가협회	18	김재희	22
5	전진우	29	전진우	17	도진기	20
6	조동신	28	한이	17	김탁환	19
7	황세연	26	김재희	17	서미애	13
8	도진기	26	정해연	17	김영탁	12
9	서미애	25	홍성호	15	전진우	12
10	한국추리작가협회	23	한새마	14	김내성	11
11	윤자영	22	송시우	14	강지영	10
12	김탁환	22	김유철	13	손선영	10
13	송시우	21	김주동	13	윤자영	10
14	이상우	19	서미애	12	이상우	10
15	조영주	18	윤자영	12	김진명	9
16	한새마	17	조영주	12	노원	9
17	한이	17	홍정기	12	장용민	9
18	홍성호, 손선영, 강지영	16	김범석	11	최혁곤	8
19			김영민	11	허수정	8
20			장우석	10	공한K, 송시우	7

먼저, P2 DB에서 가장 높은 빈도수를 차지하는 인물은 김성중이다. 이는 1970년대 등장한 김성중의 영향력이 2025년 검색 시점까지도 한국 추리 소설 지형의 한 축을 견고하게 지탱하고 있음을 의미한다. 1974년 <최후의 증인>으로 등장한 김성중은 해방 이후 침체되었던 장르적 명맥을 재건하며 김내성 이후 최대의 장르적 성취를 이루었다.²⁴⁾ 특히 『일간스포츠』²⁵⁾ 등

대중 매체를 통한 다년간 연재 활동과 뒤이은 단행본의 흥행은 한국 추리소설이 비로소 독자적인 상업성을 갖춘 '장르 문학'으로 안착하는 계기가 되었고, 이후 1983년 '한국추리작가협회'의 발족과 함께 신인 작가들이 등장하면서 추리소설이 부흥이 본격화되었다.²⁶⁾

한편, 현대 작가군에서는 개인 단행본 발간과 앤솔리지 참여가 병행되는 양상이 눈에 띈다. 정명섭(앤솔리지 25종, 단행본 25종)을 비롯해 정해연, 김재희, 전건우 등은 개인 창작과 공동 작업을 유기적으로 수행하며 장르 생태계를 확장하고 있다. 반면 조동신(총 28종 중 앤솔리지 23종)처럼 공동 작품집을 주된 활동 기반으로 삼는 사례도 발견된다.

여기서 주목할 점은 앤솔리지(135종)가 지니는 데이터의 의미다. 앤솔리지는 단순한 작품 모음집을 넘어, 각 작품집에 참여한 작가들 사이의 협업 관계와 네트워크를 보여주는 '관계망 데이터'로 이해해 볼 수 있다. 이는 주로 1인 저자 체제였던 과거의 출판 방식에서 벗어나, 현대 미스터리 장르가 작가 간의 네트워크를 통해 질적·양적 풍요를 도모하는 '공동체적 생산 모델'로 변화했음을 시사한다.²⁷⁾ 따라서 앤솔리지 내 작가군에 대한 정밀한 분석은 향후 한국 미스터리 장르의 작가 네트워크를 파악하는 데 필수 과제가 될 것이다.

24) 안혜연, 「1970년대 후반~1980년대 한국 추리소설 붐과 그 사회적 맥락」, 220-222쪽 참고.

25) 1970년대 중후반 김성종의 활약을 확인하기 위해서는 <일간스포츠>에 연재된 글들의 목록을 확인하는 것이 중요하다. 다만 본 연구에서는 이 자료를 확보하지 못했으므로 그 한계를 인정하며 추후 연구의 보완으로 이 과제를 남겨두고자 한다.

26) 안혜연, 앞의 글, 208쪽.

27) 다만 단독 단행본과 앤솔리지를 동일하게 계측할 때 발생하는 데이터 왜곡 문제는 개선이 필요하다. 예를 들어 정명섭의 사례처럼 기획 중심의 다작 활동은 실제 창작 밀도에 비해 작가 빈도수 데이터가 과대 계측될 소지가 있다. 따라서 향후 DB 설계 시에는 저작 형식에 따른 가중치를 차등 부여하고, 작가와 작품 간의 매칭 방식을 정교화하여 '허수'를 보정해야 한다. 이와 관련해서는 4장에서 후술하겠다.

2) 발행처(Publication) 분석

총 발행처는 368곳이며, 이중 단행본(Book)이 357곳, 앤솔로지는 40곳이다. 단행본과 앤솔로지를 모두 발행하는 출판사²⁸⁾는 29곳이다.

순위	발행처	빈도	순위	발행처	빈도
1	부크크(bookk)	166	11	몽실박스	27
2	고즈닉이엔티	70	12	책과나무	26
3	황금가지	63	13	아프로스미디어	23
4	남도출판사	42	14	네오픽션	23
5	엘릭시르	40	15	명지사	20
6	안전가옥	38	16	손안의책	18
7	팩토리나인	38	17	들녘	18
8	시공사	32	18	해냄출판사	17
9	나비클럽	31	19	민음사	17
10	북오션	28			

이 발행처 데이터 통계는 한국 미스터리 출판 시장의 구조적 변화를 포착할 수 있게 해준다. 주목할 만한 특징은 자가출판(Self-Publishing) 및 주문형 출판(POD) 플랫폼인 ‘부크크(bookk)’가 166종으로 1위를 차지했다는 점이다. 부크크 출판사는 제작, 납본, 출판, 유통까지 담당하는 “국내 유일의 자가출판(Self-Publishing, POD) 플랫폼”이다.²⁹⁾ 이는 기성

28) P1에서는 단행본(Book)과 단편 및 연재물(Article)의 발행처가 각기 달랐으나, P2에서는 단행본(Book)과 앤솔로지(Anthology) 모두 ‘책’의 형태로 출간되기 때문에 공통되는 부분이 생긴다.

29) POD란 ‘Publish on Demand’의 약자로, 책을 주문을 받은 만큼 소량 제작하여 보내주는 방식의 ‘주문형 출판’을 의미한다. (부크크 홈페이지, ‘부크크 소개’ 파일 참고. <https://file.bookk.co.kr/resources/primaryFiles/pressKitFile/primaryFile/1713428365801-2249.pdf>)

출판사의 '선별'을 거치지 않고도 작가가 직접 독자와 만날 수 있는 플랫폼 기반의 출판 생태계가 미스터리 장르 내에 구축되었음을 의미한다. 제목 수정 단계에서 작품 통계를 위해 제외된 [POD] 표기 도서 외에도 134종에 달하는 발행 규모는, 장르 문학의 생산 주체가 개인 창작자로 변화하고 있음을 의미한다.³⁰⁾ 출판 시장 역시 장르 전문성이 강화되는 양상을 보인다. 고즈닉이엔티(2위), 황금가지(3위), 엘릭시르(5위), 안전가옥(6위) 등 상위권 발행처들은 대부분 미스터리를 비롯한 장르 문학을 전문적으로 기획·개발하는 전문 기업 혹은 장르 특화 브랜드의 성격을 띤다.³¹⁾ 특히 민음사의 '황금가지'와 문학동네의 '엘릭시르'처럼 대형 출판사에서 별도의 장르 브랜드를 통해 국내외 미스터리 문학을 소개하고 신진 작가를 발굴하는 전략은 한국 미스터리 장르가 문학 시장의 변방에서 핵심적인 콘텐츠 공급원으로 격상되었음을 뜻한다. 한편, '남도출판사'(4위)³²⁾의 순위

30) 기존의 한국문학 출판은 공신력 있는 기관의 신춘문예를 통해 등단한 소위 '기성' 작가나 문학 전문 출판사의 엄격한 선별을 거쳐 진행되는 전통적 방식에 의존했다면, 현대 미스터리 장르는 그 생산 경로가 다변화되었다. 특히 자가출판(POD) 플랫폼의 부상은 신인 작가의 진입 장벽을 낮추었으며, 웹 기반 플랫폼에서 선 연재 및 완결 후 그 인기를 바탕으로 단행본을 출간하는 '수익 모델의 역전 현상'도 보편화되었다. 따라서 본 DB에서 나타나는 발행처의 편중과 플랫폼의 강세는 단순히 수치상의 기록을 넘어, 기술 환경 변화에 따른 장르 문학의 생산 및 유통 구조의 근본적인 패러다임 전환을 보여주는 지표로서 활용될 수 있을 것이다.

31) 고즈닉이엔티는 스릴러, 미스터리, K-Faction, 로맨스, SF, 웹소설 장르를 다루며 "IP 전문 개발사"면서 "엔터테인먼트 콘텐츠 기업"으로 소개할 수 있으며, (고즈닉이엔티 홈페이지 'ABOUT US' 참고(<https://www.gozknoock.com/posts/contact/#n>)) '황금가지'는 민음사 출판그룹으로 "SF, 추리, 판타지, 호러 등 주류 문학의 변방에 있던 젊은 문학"을 주로 다루는 출판사이다. (황금가지 홈페이지 '소개' 참고(<https://goldenbough.minumsa.com/about/>)) '엘릭시르'는 출판그룹 문학동네의 장르문학 브랜드로 동서고금의 여러 국내외 미스터리를 소개 및 격월간 미스터리 매거진 『미스터리아』를 발간하는 출판사로 "미스터리, 판타지, SF, 호러 등 다양한 해외 소설을 소개하는 동시에 한국 작가를 적극적으로 발굴"하고 있다고 밝혔다. (문학동네 홈페이지 엘릭시르 소개 참고(<https://munhak.com/munhak/etc/all#brand>))

32) '남도출판사'는 김성종과 그의 동생 김인종이 운영하는 출판사이면서 실제로 김성종

는 작가 김성종의 영향력과 긴밀히 연결된다. 김성종 작가 개인이 운영하는 이 출판사는 한 작가의 전속 출판사로서 기능하고 있음이 확인된다.

순위	1	2	3	4	5	6	7	8	9	10
앤솔로지	나비클럽	고즈닉이엔티	황금가지	청어람	부크크bookk	안전가옥	태동출판사	산더슬	화남출판사	북오션
빈도	24	11	10	8	7	7	6	5	5	5
단독출간	부크크bookk	고즈닉이엔티	황금가지	남도출판사	엘릭시르	팩토리나인	사공사	안전가옥	책과나무	북오션
빈도	159	59	53	42	39	38	31	31	26	23

출판사별 세부 발간 양상을 분석하면 단독 출간과 앤솔리지 중심의 차이가 확인된다. 앤솔리지 부문에서 가장 높은 빈도를 보이는 ‘나비클럽’이 대표적이다. 나비클럽은 <한국추리문학상 황금펜상 수상작품집>과 <괴이한 미스터리> 시리즈를 기획하고 잡지 『계간 미스터리』를 발행하며, 다수 작가의 단편을 발굴하고 소개하는 역할을 수행한다. 반면 단독 출간은 이은영의 『우울의 중점』(2021, 2022), 홍선주의 『푸른 수염의 방』(2023), 김세화의 『타오』(2024, 2025), 무경의 『마당 흑조는 곤란한 이야기를 청한다』(2024), 『부디 당신이 무사히 타락하기를』(2025) 등 적은 종수를 보인다. 이런 행보는 한국 추리소설사의 가치 있는 작품들을 지속적으로 아카이빙하고 기성 작가들의 업적을 재조명하는 데에 중점을 두는 것으로 해석된다. 특히 선집을 통해 기성 작가와 신진 작가 간 만남의 장을 마련하고 장르 문학에 대한 전문 비평의 영역을 확장함으로써, 장르적 정통성을 확립하는 일종의 ‘플랫폼’ 역할을 수행하는 것으로 분석해볼 수 있다.

의 책을 주로 발간하고 있으므로, 김성종이 P2 DB에서 가장 높은 빈도수를 보이는 작가라는 점에서 연관성을 찾아볼 수 있다.

이어 연도별 출간 횟수를 살펴보면 한국 미스터리 장르의 가파른 성장세가 확인된다. 본 DB가 온라인 서점의 데이터 보존 시기(알라딘 1976년~, 교보문고 2010년~)에 따른 제약을 지남에도 불구하고, 2019년 이후 매해 100권 이상의 출간 기록이 유지되고 있다는 점은 인상적이다. 특히 2023년은 단독 출간 158회, 앤솔리지 19회로 총 177권이 집계되며 최고치를 기록하였다. 이러한 양적 팽창은 2010년대 후반부터 본격화된 장르 문학의 시장성 강화와 더불어, 코로나19 팬데믹 시기의 독서 수요 증가, 웹소설 플랫폼과의 연계 활성화 등이 맞물린 결과로 해석된다.

이로써 확인되는 P2 DB에서의 발행처 특징은 플랫폼을 통한 자가 출판의 활성화와 장르 전문 출판사의 증가이다. 특히 2023년을 전후한 시기는 출판 주체의 다변화와 작품 스타일의 확장이라는 양적·질적 전환점을 맞이한 것으로 평가할 수 있다. 이는 한국 미스터리 문학이 과거의 지엽적인 위치를 벗어나 현대 대중 서사의 주요 장르로 부상하고 있다고도 해석할 수 있을 것이다. 그러나 이러한 양적 팽창이 기존 순문학의 주변부에 머물렀던 장르적 위상을 근본적으로 재정립하는 실질적 성장인지는 장기적인 관점에서 검토가 필요하다. 이는 향후 순문학 데이터베이스 및 타 장르 소설 데이터베이스와의 비교 분석이 병행될 때 명확히 규명될 수 있는 과제이기도 하다. 그러나 분명한 점은 파편화되어 있던 한국 추리소설의 서지 정보를 체계적으로 수집하고 정량화하는 시도 그 자체가 미스터리 장르 연구의 자생력을 확보하는 출발점이 될 수 있다는 것이다.

이로써 기초 데이터베이스 분석을 통해 시기별 작가와 발행처, 그리고 네트워크를 시각화하면서 추리 소설 장르의 특징을 입체적으로 조망해보았다. 다만, 현재의 서지 목록 중심 데이터는 1960년대부터 80년대 사이의 자료 공백과 데이터 중복 등의 한계를 지니고 있다. 이를 보완하기 위해서는 각 메타 데이터 항목별 통계 수준을 넘어 인문학적 문제의식에 따라 연

구자가 자유롭게 활용하고, 또 쿼리할 수 있는 표준화된 데이터가 필요하다. 특히 방대한 자료 수집과 저작권 문제 등 현실적인 제약을 고려할 때, 텍스트 분석에 앞서 확보된 2,200여 종의 서지 데이터를 정교화하는 작업이 우선되어야 할 것이다. 따라서 데이터의 무결성을 확보하고 쿼리 기능 수행이 가능한 분석 환경을 조성하기 위해, 현재의 평면적인 데이터 구조를 관계형 데이터베이스(RDB) 모델로 전환할 필요가 있다. 나아가 이러한 기초 데이터베이스가 완성된 후에는 미스터리 장르의 특수성을 반영한 '라벨링' 작업의 방향도 제안하고자 한다. 이는 추리·스릴러·사회파 등 각 하위 장르가 작품 내에서 차지하는 비중을 정량적 비율로 환산하여 데이터화하는 과정을 포함한다. 이에 4장에서는 이러한 문제의식을 바탕으로 구체적인 RDB 설계 방안과 하위 라벨링 모델링 방향을 제안해보도록 하겠다.

4. 관계형 데이터베이스(RDB) 모델링과 장르 라벨링 방법 제안

한국 추리소설 서지 자료는 개별 연구자가 출간한 책 혹은 학술논문이나, 장르 전문 잡지, 국립중앙도서관, 온라인 서점 웹사이트 등에 흩어져 있어 통합 작업이 필요하다. 본 연구에서 구축한 1차 DB는 이러한 자료들을 수집·정리했다는 점에서 의의가 있지만, 평면적인 구조로 이루어져 있는 한계가 있다. 연재물이나 선집에서 작가별 기여도를 분별하거나, 작가-작품-발행처 등의 네트워크를 파악하기에 어려움이 있으며, 각 데이터 출처 간 표기 방식 차이로 같은 작품임에도 여러 번 중복되는 문제도 있다. 또한 위의 유관자료 외에도 웹툰, 웹소설, 독립출판플랫폼 등 대로운 출판 생태계 작품들도 포괄하기 위해, 데이터 표준화 작업을 진행하여 확장 가

능성을 만들고, 항목 간 관계를 체계적으로 관리할 수 있는 구조가 필요하다. 이에 본 연구는 1차 DB를 관계형 데이터베이스 모델로 전환이 필요하다고 판단했다.

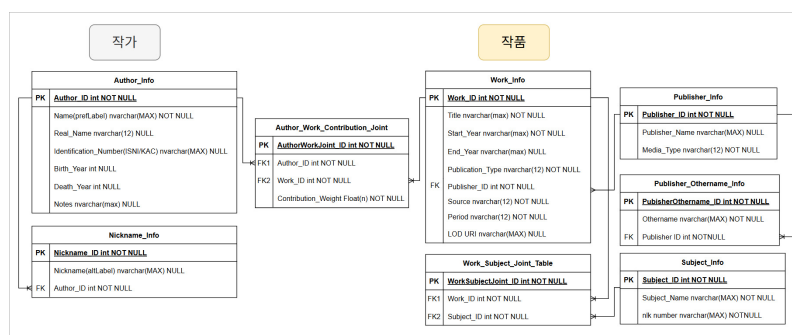
관계형 데이터베이스(RDB)는 데이터 무결성 확보에 도움이 된다.³³⁾ 예를 들면 필명을 사용하는 작가의 경우 이칭 분리를 통해 중복을 막고, 현대 데이터에 다수 포함된 'A7', '@렌' 같은 온라인 ID 형태의 작가명에는 고유 ID를 부여하고, <한국추리문학상 수상작품집>과 같은 앤솔로지 형태의 공동 저술 작품들에는 작가 가중치를 차등 부여할 수 있도록 하는 데이터 구조 설계도 가능하다. 또한 작가, 작품, 장르, 출판사, 발표 연도 등의 개별 항목을 독립된 테이블로 구성하고, 이들 간의 관계를 명확히 정의함으로써 복잡한 데이터 구조를 체계적으로 관리할 수 있다. 이러한 점에서 RDB는 미스터리 장르의 기본 데이터베이스 형식에 적합해 보인다.

이에 본 연구에서는 1차 DB를 RDB로 전환하는 모델링을 진행하였다.³⁴⁾ 먼저 2차 DB인 RDB의 ERD(Entity-Relationship Diagram, 개

33) 필자는 줄고에서 한국 근대 문인 데이터베이스를 관계형 데이터베이스(RDB)로 구축하는 작업을 진행한 바 있다. 한국 근대문학 연구에 있어서 '문인'에 관한 정확한 정보가 통합되어 있지 않고 산발적으로 흩어져 있는 문제를 해결하기 위한 시도였다. 실제로 근대 문인 정보는 필명, 동명이인, 다양한 표기 방식(국문·한자·일본어·영어·특수 기호 등) 때문에 통일이 어렵고, 신뢰할 만한 웹사이트에서도 생몰연도나 경력 정보가 상이하며 출처 확인과 수정이 곤란하다. 이에 따라 해당 연구는 명확한 출처(권영민, 『한국근대문인대사전』, 아세아문화사, 1990)에 기반한 문인 정보 데이터베이스를 설계·구축하여 기계가독형 표준 데이터를 마련하고 기존 공공자료(ISNI, LOD 등)와의 연계를 목표로 했던 것이다.(전세진·유건수·정성훈, 「한국근대문인 데이터베이스 구축 방법 탐색 - 관계형 데이터베이스 기반 설계와 그 한계를 중심으로」, 『어문연구』 제124호, 어문연구학회, 2025, 206-207쪽.) 미스터리 장르에서는 위의 경우와 같이 명확한 출처가 될 수 있는 유관 자료로서의 '사전류'가 아직 없고, 하위 장르가 다양하기 때문에 데이터베이스 구축이 더 까다롭게 진행될 수밖에 없는 상황임을 밝힌다.

34) 1차 DB를 2차 DB로 전환하는 작업이므로 1차 DB의 서지 정보를 그대로 사용하였다. 국립중앙도서관의 서지 정보까지 종합한 결과 1295명의 작가와 약 2793종의 작품, 551곳의 발행처가 확인되었다. 이 대상들의 LOD를 조사하지는 못하고, 작가 정

체-관계 다이어그램)와 제시하고, 상세 내용을 설명하도록 하겠다.



〈그림 1〉 한국 미스터리장르 - 추리소설 관계형데이터베이스(RDB)의 ERD

먼저 작가(Author)와 작품(Work)을 중심으로 핵심 테이블로 구성하였다. 작가 테이블에는 고유 식별값을 부여하여 앞서 지적된 온라인 ID 형태의 가명이나 중복 명칭, 이칭 등으로 인한 데이터 중복을 차단한다. 이를 위해 이칭(Nickname) 테이블을 만들어 대표명과 분리하여 중복을 막았다. 예를 들어 3장 1절에서 언급한 ‘꿈길’의 작가는 ‘노자영’이므로, 이를 작가 이칭(Nickname)으로 분류하면 동일 작가가 중복되는 것을 막을 수 있다.

2차 DB로 재편하며 작품 테이블에는 자료 4)의 국립중앙도서관의 서지 정보도 추가하였다. 여기에서의 문제는 자료 3)의 내용과 동일한 작품들이 발견된다는 점이고, 데이터 중복 문제를 야기할 수 있다는 점이다. Source (알라딘, 교보문고, 국립중앙도서관 등)에 따라 동일 작품에 대한 중복이 발생하는 이 문제는 LOD URL(KMO 번호)를 연계하는 설계를 통해 해결하

보는 1차로 위의 논문(「한국근대문인 데이터베이스 구축 방법 탐색」, 2025)의 자료에서 수집하였고, 현대 작가 및 작품들은 일부 샘플 작업을 진행하였다. 상세한 내용은 명세서를 확인하길 바란다. RDB 모델링 DB와 ERD, 명세서도 상기 github 저장소에서 확인할 수 있다.

고자 하였다. 또한 ‘작품(Work)’ 엔티티와 개별 출판 및 출처 정보를 관리하는 ‘출판사(Publisher_Info)’ 엔티티를 분리하고, Publisher_ID를 외래 키(FK)로 연결하는 관계형 모델을 구축하였다. 이를 통해 동일한 LOD를 가진 데이터들을 확인할 수 있게 하되, 동시에 Source 항목을 통해 각 서지 사항의 수집 경로를 명시함으로써 데이터의 계보와 신뢰도를 확보하고자 하였다. 또한 작가 테이블과의 분리를 통해 앤솔러지와 같이 한 작품에 여러 명의 작가가 있는 문제를 해결하였다.

이는 ‘작가-작품-기여도(Author_Work_Contribution)’라는 조인 테이블을 통해 살펴볼 수 있다. 이 테이블은 작가 ID, 작품 ID와 더불어 ‘참여 형태(단독 저술, 앤솔러지 참여 등)’를 ‘기여도가중치(Contribution_Weight)’ 라는 속성을 통해 작가 ‘허수’ 문제를 해결하려는 시도였다. 이는 단독 창작물에는 가중치 1.0을 부여하고, 5인이 참여한 앤솔러지 수록작에는 개별 작가에게 0.2의 기여도를 할당하는 방식이다. 아래 테이블의 샘플을 통해 알 수 있듯, 1번 130번은 단독 작가의 단행본이고, 3번은 2명의 작가가 참여한 작품, 128번은 44명의 작가의 작품이 수록된 앤솔러지임을 확인할 수 있다.

AuthorWorkJoint_ID	Author_ID	Work_ID	Contribution_Weight
1	1	2129	1
4	3	1914	0.167
69	26	1094	0.5
128	44	2339	0.023
131	47	1320	0.2
386	123	826	0.333

이러한 모델링은 특정 작가에게 발생할 수 있는 데이터 쓸림 현상을 보

정하며, 특정 시기나 매체에서 어떤 작가군이 실질적인 장르적 영향력을 발휘했는지를 더욱 객관적으로 확인할 수 있게 한다. 이는 3장에서 발생했던 특정 작가에 대한 데이터 과대 계측 문제를 기술적으로 해결하는 방향이 될 것이다.

발행처(Publisher) 테이블에서도 '매체 타입(Media_Type)' 속성을 추가하여, 출판사와 잡지, 신문사를 나누어 어떤 매체 타입이 추리 소설 장르에서 영향력이 있는지 확인할 수 있게 하였다. 나아가 이를 작품 시트에 외래키로 결합시킨 후, 작가-작품-기여도 조인 테이블을 통해 작가와 발행처 간 관계도 연결되도록 하였다. 이러한 구조적 전환은 향후 복잡한 통계 추출과 작가, 작품, 발행처 간 네트워크 분석이 더 신뢰도 있게 진행될 수 있도록 하였다.

본 연구에서 구축한 관계형 데이터베이스(RDB)의 활용 가치를 극대화하기 위해서는 서지 정보의 구조화를 넘어, 미스터리 장르의 복합적인 내부 속성을 정밀하게 파악하기 위한 멀티 라벨링(Multi-dimensional Labeling) 작업이 필요하다. 미스터리 장르는 추리, 탐정, 스릴러, 범죄, 서스펜스 등 다양한 하위 유형이 혼재되어 있어 명확한 경계 설정이 어려운 만큼, 이러한 장르적 모호성을 분석의 전제로 삼아 다중 분류 체계를 효과적으로 처리할 수 있는 설계가 요구된다.

미스터리 장르의 분류와 통합 상당히 까다로운 작업이 될 것으로 생각된다. 그럼에도 불구하고 각 하위 장르를 대표하는 텍스트 선정 및 라벨링을 통해 고유한 장르적 요소를 추출하고 체계적인 분류 기준을 마련하는 과정이 필요하다. 이를 위한 방법으로 대표 텍스트 선정 및 라벨링 작업을 통해 각 하위 장르를 대표하는 작품들을 분석하고, 고유한 장르적 요소를 추출하여 체계적인 분류 기준을 마련 및 적용하는 것을 고려할 수 있다.

현재 알라딘과 교보문고 등 온라인 서점에서 수집한 서지 정보를 분석

하면, 각 플랫폼마다 상이한 분류 체계를 사용하고 있음을 확인된다. 이 한계를 극복하기 위해, 플랫폼 간 분류 기준을 역추적하거나 텍스트 마이닝 및 기계학습 모델을 활용한 자동 분류 도구의 개발을 고려할 수 있다. 이는 단일 장르가 아닌 복합적 장르 구성을 정량적으로 파악하게 하며, 미스터리 장르의 모호한 경계를 오히려 분석의 단위로 활용하는 다차원적 태깅을 가능하게 한다.

구체적인 방법으로는 알라딘과 교보문고에서 ‘미스터리 장르’로 수집한 데이터를 비교하는 것이다. 두 플랫폼이 동일한 작품을 서로 다른 장르로 분류하는 경우가 빈번하다. 이러한 불일치를 해소하기 위한 방안으로, ‘장르소설’ 전체 데이터를 수집한 후 두 플랫폼이 공통적으로 적용하는 분류 기준을 역추적하여 비교에 적합한 대표 작품을 선정하는 것이다. 그리고 텍스트 마이닝 기법을 활용하여 장르별 특징 요소를 추출하고, 이를 기반으로 한 자동 분류 도구를 개발하는 방법이다. 추리, 탐정, 스릴러, 범죄, 서스펜스 등 각 장르의 핵심 요소—서사 구조, 어휘 패턴, 주제 등—를 기계 학습 모델에 학습시켜, 새로운 소설 텍스트를 입력했을 때 각 장르에 속할 확률을 수치화하여 제시하는 방식이다. 이러한 텍스트 분류기는 단일 장르가 아닌 복합적 장르 구성을 정량적으로 파악할 수 있게 하며, 미스터리 장르의 모호한 경계를 오히려 분석의 단위로 활용하는 다차원적 장르 태깅을 가능하게 할 것이다.

종합하자면 본 연구에서 수행한 RDB 모델링은 파편화된 한국 추리소설의 서지 데이터를 통합하고, LOD 연계 식별자 속성, 기여도 가중치 등을 통해 엄밀성을 확보해보려는 시도였다. 나아가 단순한 서지 정보를 구조화하는 것을 넘어 멀티 라벨링 작업을 제안함으로써, 기존의 미스터리 장르 문학 연구에서 계속해서 제기되어 왔던 문제들을 해결할 수 있는 방법들을 제시해보고자 하였다.

5. 결론

본 연구는 1900년대 이후 한국 추리 소설을 중심으로 한국 미스터리 장르 문학의 제문제를 정리 및 해소할 수 있는 방안들을 고찰하며, 데이터 기반의 방법론을 모색하고자 하였다. 이를 위해 근대부터 현대에 이르는 유관 자료 4종을 검토하여 2,200여 종의 추리소설 서지 정보를 수집하고 기초 데이터베이스(1차 DB)를 구축하였다. 이를 장르의 형성 및 전개기(P1)와 확장 및 다변화기(P2)로 시기를 구분하여 작가, 작품, 발행처 및 네트워크 분석을 실시하였다. 그 결과 P1에서는 추리소설이 매체 의존적인 구조 속에서 주로 연재본 형식으로 발표되었고, 문학사에서 주목받지 못했던 다양한 작가들을 발견하였다. P2에 이르러서는 자가 출판 플랫폼의 활성화와 장르 전문 출판사의 등장으로 출판 주체의 다양화를 통해 질적·양적 전환이 있었음을 확인하였다. 특히 '김성중'이라는 작가의 우세와 앤솔러지를 매개로 한 작가군 네트워크의 양상은 한국 미스터리 장르가 상업성을 갖춰나가고 있다는 현상을 확인하게 해주는 지표로 작용했다.

이러한 장르적 팽창과 서지 데이터의 복잡성을 체계적으로 수용하기 위해, 본 연구는 1차 DB의 평면적 데이터 구조를 관계형 데이터베이스(RDB) 모델로 재편찬하여 2차 DB를 설계하였다. 1차 DB에서 발견된 데이터 중복과 식별자 문제를 해결하기 위해 국가전자성명제어번호(KAC) 및 LOD(KMO 번호)를 연계하였으며, 이를 통해 저자 식별의 객관성과 데이터의 무결성을 확보하였다. 또한 앤솔러지 등 공동 저작물에 대한 작가별 기여도 가중치 설계를 도입하여 통계적 왜곡을 최소화하고, 파편화된 정보를 디지털 인문학적 관점에서 통합 관리할 수 있는 구조를 마련하였다. 나아가 2차 DB를 활용할 수 있는 방법으로, 미스터리 하위 장르의 혼

재 양상을 정밀하게 규명하기 위한 다차원 라벨링 방법을 제안하였다.

이 연구는 파편화된 한국 미스터리 서지 정보를 디지털 인문학적 관점에서 통합하고, 추후 장르 문학 연구의 토대 마련과 확장을 도모하고자 하였다. 비록 일부 시기의 자료 공백과 2차 DB(RDB) 모델링이 샘플 작업으로 진행되었으나, 그동안 산재되어 있던 한국 미스터리 서지 정보를 통합하고 분석 가능한 형태로 구조화했다는 점에서 의의를 찾아볼 수 있다. 본 연구가 제시한 데이터베이스 설계 방법론은 향후 연구자들이 데이터를 확장하고 정밀화하는 데 실질적인 참고점이 될 수 있기를 기대한다.

참고문헌

1. 기본자료

- 권영민, 『한국현대문학사』, 민음사, 2002.
김윤식·정호웅, 『한국소설사』, 문학동네, 2000.
조동일, 『한국문학통사 5(근대문학 제1기)』, 지식산업사, 2005.
『별건곤』, 『실화』, 『아리랑』, 『야담』, 『야담과 실화』, 『미스터리』

2. 논문과 단행본

- 공임순, 「1950년대 『야담과 실화』 필화 사건과 두 하위 장르의 선정성 논란」, 『대중서사연구』 제31권 2호, 대중서사학회, 2025, 167-204쪽.
곽승숙, 「방인근의 탐정소설 연구」, 『Journal of Korean Culture』 제34호, 고려대학교 한국언어문화기술혁신연구소, 2016, 69-98쪽.
김영성, 「한국 추리서사의 서사성과 대중성에 관한 연구 (1)- 추리서사의 사적(史的) 개념과 적용 범주에 대하여 -」, 『한국언어문화』 제29호, 한국언어문화학회, 2006, 183-208쪽.
_____, 「한국 추리서사의 서사성과 대중성에 관한 연구」 연구요약문 및 결과보고서, 한국연구재단, 2005/2006. (2026.01.06. 인용.)
_____, 「한국 현대소설의 추리소설적 서사구조 연구」, 한양대학교 박사논문, 2003.
김영애, 「『최후의 승리』 판본 연구」, 『현대문학의 연구』 제54호, 한국문학연구학회, 2014, 255-284쪽.
김효진, 「한국 근대 추리소설 연구」, 강원대학교 박사논문, 2014.
대중서사장르연구회, 『대중서사장르의 모든 것 3: 추리물』, 이론과 실천, 2011.
박광규, 「1950년대 한국 추리소설사의 조각 맞추기」, 『미스터리』 제19호, 2018.07.
박유희, 「한국 추리서사에 나타난 ‘탐정’ 표상—한국 추리서사의 역사와 이론을 위한 시론—」, 『한민족문화연구』 제31호, 한민족문화학회, 2009, 397-434쪽.
박인성, 『이것은 유해한 장르다』, 나비클럽, 2024.
박진영, 『탐정의 탄생』, 소명출판, 2018.

- 안혜연, 「1970~80년대 한국 추리소설 연구」, 성균관대학교 박사학위논문, 2022.
- _____, 「1970년대 후반~1980년대 한국 추리소설 붐과 그 사회적 맥락」, 『인문과학』 제85호, 성균관대학교 인문학연구원, 2022, 193-234쪽.
- 연남경, 「김유정 소설의 추리 서사적 기법 연구」, 『한중인문학연구』 제34호, 한중인문학회, 2011, 55-79쪽.
- 오혜진, 「1930년대 한국 추리소설 연구」, 중앙대학교 박사학위논문, 2008.
- _____, 「1950~90년대까지 추리소설의 전개 양상」, 『어문논집』 제44호, 민족어문학회, 2010, 299-320쪽.
- _____, 『대중, 비속한 취미 '추리'에 빠지다』, 소명출판, 2013.
- 윤영천, 『미스터리 가이드북』, 한스미디어, 2021.
- 이범봉, 「탐정물, 1950년대 사회 문화의 속어」, 『미스테리아』 19, 2018.07.
- 이주라, 「근대 초기 번역·번안 추리소설의 수용양상 연구」, 『어문논집』 제61호, 민족어문학회, 2010, 81-109쪽.
- 이현지, 「한국 근대 추리소설 연구」, 단국대학교 박사학위논문, 2013.
- 전세진·유진수·정성훈, 「한국근대문인 데이터베이스 구축 방법 탐색 - 관계형 데이터베이스 기반 설계와 그 한계를 중심으로」, 『어문연구』 제124호, 어문연구학회, 2025, 191-230쪽.
- 정종현, 「해방전후 김내성 스파이-탐정 서사의 연속과 비연속-〈태풍〉과 《청춘극장》을 중심으로-」, 『현대문학의 연구』 제42호, 한국문학연구학회, 2010, 457-508쪽.
- 정혜영, 「김내성과 탐정문학」, 『한국현대문학연구』 제20호, 한국현대문학회, 2006, 405-433쪽.
- 조은숙, 「탐정소설, 소년과 모험을 떠나다-1920년대 방정환 소년탐정소설의 문학사적 위치와 의의」, 『우리어문연구』 제38호, 우리어문학회, 2010, 615-645쪽.
- 진설아, 「90년대 추리소설의 장르적 방향성과 한계 연구」, 『문화와 융합』 제44권 1호, 한국문화융합학회, 2022, 351-366쪽.
- 차선일, 「한국 근대 탐정소설 연구」, 경희대학교 박사학위논문, 2012.
- 최애순, 「〈학원〉의 해외 추리·과학소설의 수용 및 장르 분화 과정」, 『대중서사연구』 제21권 3호, 대중서사학회, 2015, 275-320쪽.

_____, 「1930년대 探偵의 의미 규명과 探偵小説의 특성 연구」, 『동양학』 제42호, 단국대학교 동양학연구원, 2007, 23-42쪽.

_____, 「1960년대 서스펜스 스릴러의 유입과 신진 추리소설가의 등장—허문녕, 천불란, 천세욱의 단행본을 중심으로—」, 『현대소설연구』 제89호, 한국현대소설학회, 2023, 247-285쪽.

Ekaterina Beglova, 「김성중 추리소설 연구-80년대 장편을 중심으로」, 서울대학교 석사논문, 2020.

Abstract

A Preliminary Study on the Database Design for Mystery Literature - Focusing on Korean Detective Novels

Jeon, Se-Jin(Seoul National University of Science and Technology)

This study proposes a methodology for constructing and refining databases of Korean detective novels through the collection and quantification of bibliographic materials. Despite the rapid expansion of Korean mystery literature since the 21st century, systematic bibliographic information needed to understand its development remains insufficient. To address this, the study collected approximately 2,200 bibliographic records spanning from the early 1900s to the present, constructed a primary database, and divided them into two periods, and conducts quantitative analysis.

Based on these findings, the study develops a relational database(RDB) model to process complex bibliographic information and ensure data integrity, and proposes a multidimensional labeling system reflecting the composite attributes of subgenres. This research provides an objective diagnosis of Korean detective novel studies and establishes a methodological foundation for genre literature research, with potential for future expansion into narrative analysis through text mining.

(Keywords: Mystery Literature, Korean Detective Novel, The Database Design, Digital Humanities, Relational Database, Genre Labeling)

논문투고일 : 2026년 1월 10일

심사완료일 : 2026년 1월 27일

수정완료일 : 2026년 2월 10일

게재확정일 : 2026년 2월 14일