

조선 의학 병문(病門)의 판단 구조 데이터 모델링 연구: 의방유취·동의보감·인제지·역시만필 3 축+1 체계를 중심으로

Modeling Diagnostic Algorithms in Joseon Medicine: Designing Research Data Modules for 30 Byeongmun (病門) through a Three-Axis-Plus-One Textual Framework

전종욱 (Jeon, Jong-Wook)*

 0000-0002-2194-3361

유인태(Ryu, In-Tae)**

 0009-0004-3725-8248

목차

1. 서론
 - 1.1 기존 연구의 흐름과 공백
 - 1.2 연구 목적과 범위
2. 대상 문헌 체계: 3 축+1 의 설정 근거
 - 2.1 『의방유취』: 축적의 층위
 - 2.2 『동의보감』: 표준화의 층위
 - 2.3 『인제지』: 응용·재편의 층위
 - 2.4 『역시만필』: 임상 검증의 층위
 - 2.5 3 축+1 체계의 의미
3. RDM 설계 원칙과 데이터 모델
 - 3.1 3 대 원칙: 문헌 우선·분리·구조화
 - 3.2 병문(病門): 분석 단위
 - 3.3 L1·L2·L3 병문 내부 구조의 세 층위 도출과 정의
 - 3.4 병문 유형 분류 체계(Type A-D)의 귀납적 도출
 - 3.5 Provenance_Citation 설계와 문헌 계보 보존
 - 3.6 데이터 스키마의 뼈대
4. 편찬 데이터와 그 특징
 - 4.1 30 병문 데이터셋 개요와 DataON 등록 현황
 - 4.2 유형별 분포와 구조적 특성
 - 4.3 병문 간 관계 데이터의 패턴 분석
 - 4.4 데이터에서 귀납된 두 특성: 원인배열과 개입타이밍 조율
5. 한계점 및 보완 방향: 관련 연구와의 비교

*제 1 저자: 전북대학교 한국과학문명학연구소, 부교수, lovejnj@jbnu.ac.kr

**교신저자: 전남대학교 중어중문학과, 조교수, ryuintae@gmail.com

- 5.1 국내 한의학 디지털화 프로젝트와의 차이
- 5.2 해외 전통의학 디지털화 프로젝트와의 비교
- 5.3 본 연구의 한계와 후속 연구의 방향
6. 지식그래프 확장과 AI 연계의 전망
- 6.1 3층 지식그래프로의 확장 설계
- 6.2 AI 활용 데이터로서 가능성과 조건
- 6.3 전통 지식 데이터의 FAIR·CARE 원칙 적용
7. 결론

초록

본 논문은 조선 의학 고전문헌에 내재된 질병 판단 알고리즘을 데이터로 설계·구축하는 방법론과 그 결과를 제시한다. 『의방유취』·『동의보감』·『인제지』를 축적·표준화·응용의 3 축으로, 『역시만필』을 임상 검증 텍스트로 설정하는 ‘3 축+1’ 문헌 체계 위에서 30 개 병문(病門)의 판단 흐름을 L1(현상 인식)·L2(판단 구조)·L3(치료 실행)의 3 층위로 분절하고, 병문의 판단 기능에 따라 생태순환형(A) · 판단노드형(B) · 네트워크허브형(C) · 경계판단형(D)의 4 유형으로 분류하였다. 구축된 데이터셋은 RDM_Core 약 750 행, NETWORK 관계 약 485 건으로 구성되며, KISTI DataON 국가연구데이터플랫폼에 공식 등록되었다. 분석 결과 L2(판단 구조)가 전체의 약 45%를 차지하여 조선 의학 문헌이 판단의 논리 서술에 가장 큰 비중을 두고 있음이 확인되었고, Type C(네트워크허브형)의 평균 NETWORK 엣지 수가 다른 유형보다 뚜렷하게 높아 유형 분류의 정합성이 사후적으로 검증되었다. 그리고 데이터에서 두 가지 구조적 특성이 귀납적으로 확인되었으니, 복합 원인을 단일 요인으로 환원하지 않고 체계적으로 배열하여 판단을 구성하는 ‘원인배열의학’과 개입의 시점·순서·강도를 병의 진행 단계에 따라 신중하게 조율하는 ‘개입타이밍 조율의학’이 그것이다. 이러한 특성이 기존 해외 전통의학 디지털화 프로젝트의 접근 및 방법론과 구별됨을 논증하고, AI와의 연계 가능성을 함께 제시하였다.

주제어: 조선 의학, 병문(病門), 데이터 모델, 원인배열의학, 개입타이밍 조율의학

Abstract

This paper presents a methodology for designing and constructing data that restores the diagnostic algorithms embedded in classical Joseon medical literature, along with its results. Setting *Uibangryuchi* (醫方類聚), *Donguibogam* (東醫寶鑑), and *Injeji* (仁濟志) as a three-axis framework representing accumulation, standardization, and application, and *Yeoksimanpil* (歷試漫筆) as a clinical verification text in a "Three-Axis-Plus-One" system, the study segments the diagnostic flow of 30 *byeongmun* (病門, disease categories) into three layers: L1 (phenomenal recognition), L2 (judgment structure), and L3 (therapeutic execution). Each *byeongmun* is classified into one of four typologies based on its functional role: ecological-circulation (A), judgment-node (B), network-hub (C), and boundary-judgment (D). The dataset comprises approximately 750 rows of RDM_Core data and 485 NETWORK relations, officially registered on the KISTI DataON Platform. L2 (judgment structure) accounts for about 45% of the total, confirming that Joseon medical literature allocates the greatest proportion to articulating diagnostic logic. The mean NETWORK edge count for Type C (22.1) markedly exceeds the other types (11.2–16.2), providing post-hoc validation of the classification. Two structural characteristics were inductively identified: *Cause-Arrangement Medicine*, which systematically arranges multiple causes rather than reducing them to a single etiology, and *Intervention-Timing Medicine*, which calibrates the timing, sequence, and intensity of therapeutic intervention by disease stage. These characteristics are methodologically distinct from the nominal entity-centered approaches in existing international traditional medicine digitization projects, opening prospects for knowledge graph expansion and AI integration.

Keywords: Joseon medicine, *byeongmun* (病門), data model, Cause-Arrangement Medicine, Intervention-Timing Medicine

1. 서론

1.1 기존 연구의 흐름과 공백

조선 의학에 대한 학술 연구는 크게 다음 몇 개의 흐름으로 축적되어 왔다. 첫째는 포괄적 의학사 연구다. 역사학·사회사·문화사·제도사의 시야에서 조선 의학을 외부 맥락과 함께 해석하는 흐름으로, 의학 지식 자체보다도 그것이 형성·유통·계승되는 조건에 주목한다.¹ 둘째는 한의학의 학문적 기반을 근간으로 하는 원전·이론 연구가 하나의 큰 줄기를 형성한다. 『대한한의학원전학회지』, 『한국의사학회지』 등을 통해 축적된 연구 성과가 이 흐름의 중심에 있다.² 조선 의학의 외적 맥락을 포괄적으로 다루는 첫째 흐름과 달리, 둘째 흐름은 한의학이라는 학문 전통의 내부에 서서 지식 자체의 내적 구조를 정밀하게 독해하는 데 중점을 둔다. 셋째는 고전 의서의 정리·번역·디지털화 연구다. 한국한의학연구원이나 특허청이 운영하는 데이터베이스 기반 웹사이트가 이 범주에 속하며, 한의학 고문헌·전통 처방·약재·화합물 정보를 구조화하여 국제특허분류 체계와 연동한 국가적 인프라로 기능해 왔다.³ 이 흐름은 원전 접근성의 확보와 전통지식의 법적·제도적 보호라는 기반 구축에 기여해 왔다. 넷째는 임상 응용·신약 개발과 직접 연계하여 근거 기반의 데이터를 축적·제공하는 연구다. 경희대학교와 한국한의학연구원에서 운영하는 웹사이트가 대표적이다.⁴ 이 흐름은 전통의학 지식을 현대 임상·신약 개발의 문법으로 변환하여 활용 가능성을 높이는 데 중점을 둔다.

이 흐름들은 각각 조선 의학의 외적 맥락, 내적 이론, 텍스트 접근성, 임상·연구 응용이라는 상이한 층위에 대응하며, 조선 의학 연구의 지형을 다각도로 확장해 왔다. 한편으로 기존 연구사를 검토했을 때 아직 충분히 다루어지지 않은 영역이 있으니, 기록에 근거한 병문(病門)에 대한 심층적 연구다. 조선 의학 원전 텍스트의 내부에서 실제로 작동하는 판단의 구조(병을 어떤 순서로 인식하고, 어떤 기준으로 감별하며, 어떤 시점에 개입하고 어떤 경우에 개입을 유보하는가에 관한 사고의 흐름)를 문헌 근거 위에서 구조적으로 드러내고, 이를 재현 가능한 데이터로 구축하는 작업이 그것이라 하겠다.

관련하여 해외에서는 전통의학의 지식을 디지털 환경에서 구조화하는 작업이 빠르게 진행되고 있다. 약재·질병·증상·처방을 엔티티로 추출하여 대규모 그래프를 구축한 사례로 중국의 TCM 지식그래프⁵가 있으며, 『상한론(傷寒論)』을 대상으로 진단 사유 모델 기반의 지식그래프를 수작업으로 구축한 사례⁶도 소개되었다. 인도의 TKDL⁷은 아유르베다(आयुर्वेद, Ayurveda) 등의 전통 처방들을 독자적 분류 체계로 구조화하여 국제특허분류와 연동하였으며, 최근에는 아유르베다의 약용식물·화학성분·처방·질병을 통합한 그래프 데이터베이스 GRAYU⁸가 발표되었다. 싱가포르에서는 동남아시아 전통의학의 다언어 약재 지식을 Neo4j 기반 지식그래프로 구현한 Polyglot Asian Medicine 프로젝트⁹가 진행 중이다. 말레이 의학 필사본을 온톨로지 기반으로 구조화한 연구¹⁰도 최근 소개된 바 있다.

이러한 연구 및 프로젝트는 대체로 약재·처방·질병이라는 명사적 개체(Entity)의 추출과 관계(Relationship) 매핑에 초점을 두고 있다. 의학 고전 텍스트 내부에 내재된 판단의 절차적 흐름(현상을 인식하고, 원인을 감별하며, 개입 여부와 시점을 결정하는 순차적 구조)을 분절하여, 데이터 모델로 복원하는 접근은 아직 본격적으로 시도되지 않았다. 그리고 복수의 문헌에 걸친 인용 계보(citation lineage)를 데이터 구조 안에 내장하여 지식의 출처와 전승 경로를 추적 가능하게 만드는 설계 또한 기존 프로젝트에서는 찾아보기가 어렵다. (5장에서 상론함)

1.2 연구 목적과 범위

본 연구의 목적은 조선 의학 고전문헌에 내재된 질병 판단 알고리즘을 데이터로 설계·구축하는 방법론을 제시하고, 구축된 데이터셋의 구조적 특성을 분석·소개하는 데 있다.

이를 위해 본 연구는 『의방유취(醫方類聚)』, 『동의보감(東醫寶鑑)』, 『임원경제지(林園經濟志)·인제지(仁濟志)』를 3축으로, 『역시만필(歷試漫筆)』을 임상 검증 텍스트로 삼는 ‘3축+1’ 대상 문헌 체계를 설정하였다. 이 체계 위에서 30개 병문(病門)을 분석 단위로 삼아, 각 병문의 판단 흐름을 L1(현상 인식)·L2(판단 구조)·L3(치료 실행)의 3층위로 분절하고, 병문이 수행하는 판단 기능에 따라 4가지 유형(Type A-D)으로 분류하였다. 편찬된 30병문 데이터셋은 현재 KISTI DataON 국가연구데이터플랫폼에 공식 등록되어 있다.¹¹

본 논문의 구성은 다음과 같다. 2장에서는 ‘3축+1’ 문헌 체계의 설정 근거를 제시한다. 3장에서는 편찬한 데이터셋의 설계 원칙, L1·L2·L3 판단 층위의 도출 과정, 병문 유형 분류 체계, 데이터셋 스키마를 기술한다. 4장에서는 구축한 결과를 유형별·관계별로 분석하고, 데이터에서 귀납적으로 확인되는 구조적 특성을 논의한다. 5장에서는 국내외 관련 프로젝트와의 방법론적 비교 및 본 연구의 한계점에 대해 검토하고, 이어 6장에서는 추후 본 연구 결과물이 지식그래프(Knowledge Graph)로 확장될 가능성과 AI와 연계 가능한 학술적 전망에 대해 제시한다. 마지막 7장에서 본고의 종합 결론을 서술한다.

2. 대상 문헌 체계: 3축+1의 설정 근거

조선 의학의 판단 구조를 재현하려면, 지식이 축적되고 표준화되며 응용·재편되고 임상에서 검증되는 과정 전체를 포괄할 수 있는 문헌 체계가 필요하다. 이를 위해 본 연구는 데이터 모델의 근거를 특정한 단일 문헌에 의존하지 않고, 『의방유취(醫方類聚)』, 『동의보감(東醫寶鑑)』, 『임원경제지·인제지(仁濟志)』를 3축으로, 『역시만필(歷試漫筆)』을 임상 검증 텍스트로 설정하는 ‘3축+1’ 체계를 설정하였다. 이 장에서는 각 문헌이 어떠한 역할과 기능을 하는지, 그리고 네 문헌이 어떤 상호 관계 속에서 하나의 분석 체계

를 구성하는 지에 대해 밝힌다.

2.1 『의방유취(醫方類聚)』: 축적의 층위

15세기 세종 대에 편찬되고 성종 대 간행된 『의방유취』는 당시까지 존재한 중국과 조선의 의학 지식을 365권(266권 간행) 규모로 집성한 저작이다.¹² 이 문헌의 핵심적 의의는 방대한 분량에 있는 것이 아니라, 모든 의학 지식을 91개 병문(病門) 단위로 분류하고 배열한 편제 방식에 있다. 특정 병증에 관련된 고금의 논의를 한 자리에 모아 제시하는 이 방식은, 흩어진 지식을 단순히 수집하는 것을 넘어 그것을 판단 가능한 형태로 조직하려는 시도였다. 이 편제 방식 자체가 이미 하나의 판단 구조를 내포하고 있다고 할 수 있다.

본 연구에서 『의방유취』가 ‘축적의 층위’로 설정되는 이유는 두 가지다. 첫째, 『의방유취』의 병문 편제는 이후 『동의보감』과 『임원경제지·인제지』의 체제적 기반이 되었다. 둘째, 『의방유취』는 하나의 병증에 대해 복수의 의서가 제시하는 서로 다른 견해를 병렬적으로 수록하고 있어, 조선의학이 단일 권위에 의존하지 않고 다원적 지식 위에서 판단을 구성했음을 보여준다. 데이터 모델링 과정에서 『의방유취』는 각 병문의 지식 원천과 인용 계보(Provenance Citation)를 추적하는 기준 텍스트로 기능한다.

2.2 『동의보감(東醫寶鑑)』: 표준화의 층위

17세기 초 허준(許浚, 1539~1615)이 완성한 『동의보감』은 『의방유취』를 통해 축적된 방대한 지식을 국가 표준으로 정제한 문헌이다.¹³ 내경(內景)–외형(外形)–잡병(雜病)–탕액(湯液)–침구(鍼灸)로 이어지는 5부 체계는 편의적으로 분류된 기준이 아니라, 당시의 우주관에서 인체를 바라보는 논리이자 문제에 접근하는 순서를 구조적으로 제시한 것이다. 임상 현장에서 어디서부터 사고를 시작하고 어느 방향으로 판단을 전개해 나가야 하는지에 대한 위계가 이 체제 안에 내장되어 있다.

본 연구에서 『동의보감』이 ‘표준화의 층위’로 설정되는 이유는, 이 문헌이 『의방유취』에 수록된 다원적 지식 가운데 무엇을 선택하고 어떤 순서로 배열할 것인지를 판단하여 하나의 표준 체계로 제시했기 때문이다. 『의방유취』가 ‘어떠한 내용이 논의되었는가’를 보여준다면, 『동의보감』은 ‘그중 무엇이 표준으로 채택되었는가’를 보여준다. 데이터 모델링 및 구축 과정에서 『동의보감』은 각 병문의 판단 위계와 구조적 골격을 제공하는 중심 텍스트로 기능한다.

2.3 『임원경제지·인제지(仁濟志)』: 응용·재편의 층위

19세기 초 서유구(徐有渠, 1764-1845)가 편찬한 『임원경제지·인제지』(이하 『인제지』)는 본 연구에서 설정한 세 번째 층위를 대표한다. 『인제지』는 당시 중국의 『본초강목(本草綱目)』, 『사고전서(四庫全書)』, 『고금도서집성(古今圖書集成)』 등이 형성한 거대한 외부 지식의 자장을 정면으로 의식하면서도, 조선의 땅과 사람에 맞는 편집체제로 지식을 재구성한 결과물이다. 치료만이 아니라 ‘병 전-병 중-병 후-양생-금기-조장’을 하나의 운영 체계로 묶어 사고하는 『인제지』의 구성은, 조선 의학의 판단이 질병 치료에 국한되지 않고 삶 전체의 운영에 관한 것이었음을 보여주는 단서라 하겠다.¹⁴

본 연구에서 『인제지』가 ‘응용·재편의 층위’로 설정되는 이유는, 이 문헌이 『동의보감』의 표준 체계를 수용하되 그것을 그대로 답습하지 않고 임상적·실용적 관점에서 재편했기 때문이다. 『의방유취』가 축적을, 『동의보감』이 표준화를 대표한다면, 『인제지』는 표준화된 지식이 구체적 맥락에서 어떻게 응용·변형되는지를 보여준다. 데이터 모델링 및 구축에서 『인제지』는 『동의보감』의 판단 구조가 후대에 어떻게 계승·변용되었는지를 추적하는 비교 텍스트로 기능한다.

2.4 『역시만필(歷試漫筆)』: 임상 검증의 층위

17~18세기 어의(御醫) 이수귀(李壽龜, 1664~1743 이후)의 『역시만필』은 앞서 정리한 세 텍스트와는 그 성격이 다르다.¹⁵ 『의방유취』, 『동의보감』, 『인제지』가 규범 문헌(병을 어떻게 인식하고 판단해야 하는가에 대한 체계적 서술)이라면, 『역시만필』은 그 규범이 실제 임상에서 어떻게 작동했는지를 기록한 의안(醫案) 모음이다. 130개 의안에는 환자의 성별·신분·직업까지 기록되어 있어 조선 의학사의 중요한 사료적 가치를 지닌다.

본 연구에서 『역시만필』이 ‘3축’이 아닌 ‘+1’로 설정되는 이유는, 이 문헌의 기능이 다른 세 문헌과 구별되기 때문이다. 『역시만필』은 판단 체계를 직접적으로 제시한다기보다도, 3축 문헌이 제시하는 판단 체계가 실제 환자에게 적용될 때 어떤 양상을 보이는지를 사례로 보여준다. 규범과 실천 사이의 간극, 또는 일치를 확인할 수 있는 검증 텍스트인 것이다. 데이터 모델링 및 구축 과정에서 『역시만필』은 L1·L2·L3 층위로 분절된 판단 구조가 실제 임상 기록과 정합하는가를 교차 확인하는 데 활용된다고 하겠다.

2.5 3축+1 체계의 의미

이 체계의 설계 논리는 다음과 같다. 3축은 조선의학 지식 생산의 세 국면 즉, 다원적 축적, 국가적 표준화, 맥락적 응용을 각기 대표하며, 이 셋을 함께 살펴봄으로써 단일 문헌 하나에 종속되지 않는 복합적 시야를 확보한다. ‘+1’로 배치된 『역시만필』은 규범

문헌이 제시하는 판단의 ‘당위’를 실제 임상의 ‘실행’ 기록과 대조할 수 있게 해 준다.

이 체계에서 중요한 것은 네 문헌이 시간순으로 나열된 연대기가 아니라는 점이다. 축적→표준화→응용이라는 지식 생산의 기능적 층위에 따른 배치이며, 『역시만필』은 시기적으로는 『동의보감』과 『인제지』 사이에 위치하지만 기능적으로는 세 축 전체를 검증하는 별도의 위상을 지닌다. 본 연구의 데이터 모델이 단일 텍스트를 기반으로 한 기존 유관 프로젝트와 구별되는 첫 번째 지점이 바로 이러한 ‘기능적 층위에 따른 다수의 문헌 배치’에 있다고 하겠다.

표 1. ‘3 축+1’ 문헌 체계의 구조

문헌	시기	층위	데이터 모델에서의 기능
의방유취	15 세기 초	축적	지식 원천·인용 계보 추적의 기준
동의보감	17 세기 초	표준화	판단 위계·구조적 골격의 중심
인제지	19 세기 초	응용·재현	표준 체계의 계승·변용 추적
역시만필	17~18 세기	임상 검증	판단 구조의 실천적 교차 검증

3. RDM 설계 원칙과 데이터 모델

이 장에서는 30병문 RDM의 설계 원칙, 판단 층위의 도출 과정, 병문 유형 분류 체계, 데이터셋 스키마를 상술한다. 본 연구의 핵심적 방법론이 이 장에 집약되어 있다고 하겠다.

3.1 문헌 우선·분리·구조화

본 프로젝트의 데이터 모델링은 세 가지 기본 원칙 위에 서 있다.¹⁶

첫째, 문헌 우선 원칙이다. 모든 데이터는 원전 텍스트에서 직접 추출하며, 연구자의 해석적 개입을 최소화한다. 특정 병문에 대한 연구자의 의학적 판단이나 이론적 해석이 데이터에 선행하는 것이 아니라, 문헌의 서술 순서와 구조가 데이터 정리의 기초적 기준이 된다.

둘째, 분리 원칙이다. 한문 원문, 한글 번역문, 문헌 3축의 출처 정보, Provenance_Citation 정보를 서로 혼합하지 않고 별도의 필드로 분리하여 기록한다. 이 원칙은 원문과 해석이 뒤섞여 출처를 추적할 수 없게 되는 문제를 방지하기 위한 것이다.

셋째, 구조화 원칙이다. 문헌의 연속적 서술을 ‘판단 단위’로 분절하여 RDM 데이터 구

조에 맞게 정리한다. 여기서 판단 단위란 하나의 독립적인 인식·감별·처치 판단을 담고 있는 텍스트 구간을 말한다.

이 세 원칙을 관통하는 핵심 명제는 ‘문장이 스스로 말하게 하라’는 것이다. 데이터는 기본적으로 문헌의 구조를 반영해야 하며, 연구자의 설명이 문헌의 목소리를 대체해서는 안 된다.

3.2 병문(病門): 분석 단위

데이터의 기본 분석 단위는 개별 질병명이 아니라 ‘병문(病門)’이다. 병문은 『의방유취』 이래 조선의학 문헌이 채택해 온 지식 편제의 기본 단위로, 단일 질병이 아니라 특정 병리적 범주 아래 관련된 증상·원인·감별·처방을 묶어 놓은 의학 판단의 구조 단위라 할 수 있다.

예컨대 해수문(咳嗽門)은 ‘기침[咳嗽]’이라는 단일 증상이 아니라, 기침을 중심으로 한 호흡기 병증 전체의 인식·감별·치료 판단을 포괄한다. 담음문(痰飲門)은 ‘담’이라는 물질이 아니라, 체액 병리가 해수·구토·현훈·심계 등으로 전이되는 복합 네트워크를 판단하는 구조 단위다. 이처럼 “병문은 질병의 이름표가 아니라 판단의 장(場)”이며, 이것이 개별 질병명이 아니라 병문을 데이터의 편찬 단위로 채택한 이유다.

본 연구의 1차 대상은 30개 병문이다. 실제로 『의방유취』는 91개 병문, 『동의보감』은 약 68개 병문(내경·외형·잡병편 기준), 『인제지』는 약 40개 병문으로 파악할 수 있는데, 세 문헌의 편제 단위가 서로 다르기 때문에 단순 숫자로 선정 범위를 표현하기는 어렵다. 30병문은 세 문헌 모두에서 공통적으로 다뤄지는 병문을 우선으로 하되, 편제 차이에 따라 병문의 범위를 조정·정렬한 결과다.¹⁷ 굳이 선정 기준을 꼽자면 다음 세 가지를 들 수 있다. 첫째, 3축 문헌 모두에서 해당 병문이 독립적으로 다루어져 있어 교차 분석이 가능할 것, 둘째, 병문의 판단 구조가 단순 나열이 아닌 감별·분기·전이 등의 판단 흐름을 내포하고 있을 것, 셋째, 가능한 경우 『역시만필』 관련 의안 또는 임상 사례와 교차 검토할 수 있을 것 등이다.

3.3 L1·L2·L3: 병문 내부 구조의 세 층위 도출과 정의

각 병문의 내부 구조는 세 가지 판단 층위로 분절된다. 이 층위 구분은 사전에 이론적으로 설정한 것이 아니라, 복수의 병문 텍스트를 반복적으로 읽고 분절하는 과정에서 귀납적으로 도출된 것이다. 도출 과정을 구체적으로 서술하면 다음과 같다.

조선 의학 문헌의 병문 서술은, 문헌에 따라 분량 및 상세함에 차이가 있으나, 공통적으로 일정한 서술 순서를 따른다. 대체로 해당 병증이 무엇인지를 규정하는 서술이 먼저 오고, 그 병증의 원인과 감별에 관한 논의가 이어지며, 마지막으로 구체적인 처방과 개

입의 방법이 제시된다. 이 반복적 패턴을 여러 병문에 걸쳐 확인한 결과, 세 개의 판단 층위로 분절하는 것이 텍스트의 내재적 구조에 가장 부합한다는 결론에 도달하였다. 세 층위의 정의는 다음과 같다.

세 층위의 정의는 (가) 기능적 역할(functional role), (나) 텍스트 표지(linguistic markers), (다) 전형적 한문 어구 패턴(typical lexical patterns)의 세 차원에서 제시된다. 이 세 차원이 동시에 충족되는 정도에 따라 각 데이터 행의 층위가 배정된다. 판별 기준 자체도 사전에 이론적으로 부여된 것이 아니라, 30병문 텍스트를 반복적으로 읽고 분절하는 과정에서 어구 패턴의 규칙성이 가장 안정적으로 확인되는 지점을 사후적으로 정리한 결과다. 다음 표 2는 그 기준을 정리한 것이다.

표 2. L1·L2·L3 판별 기준: 기능적 역할·텍스트 표지·전형적 한문 어구 패턴

층위	기능적 역할	텍스트 표지	전형적 한문 어구 패턴 (사례 포함)
L1 현상 인식	병증의 범위·정체·외형을 규정·기술하는 층위. 정의(definition)와 외형 기술(description) 이 일차 기능. '이 병증이 무엇인가'를 확정한다.	명명·정의의 어구, 증상의 외형적 묘사, 발생 양상의 일반화 서술, 병증 명칭의 분절·구분	정의: 「A 謂 B」, 「A 者 … 也」, 「凡 A 曰 B」, 「夫 A …」. 외형 기술: 「其證 …」, 「形如 …」, 「狀如 …」. 사례(해수문): 「咳謂無痰而有聲 肺氣傷而不清也」(인제지, COU-M01), 「嗽謂無聲而有痰 脾濕動而爲痰也」(인제지, COU-M02).
L2 판단 구조	병의 원인을 복수로 배열하고, 감별 분기를 제시하며, 병문 내·외부로의 전이·전변 경로를 서술하는 층위. 배열·감별·전이의 서술이 일차 기능.	원인의 다원성 선언, 단일 환원의 거부, 감별 분기(虛實·陰陽·表 裏·寒熱·內外), 인과·시간적 진행, 전이·전변·귀속의 서술	다원성/환원 거부: 「A·B·C 皆能 …」, 「皆令人 …」, 「非獨 … 也」. 감별·귀속: 「屬 A 屬 B」, 「分爲 A·B·C」. 인과: 「因 A 而 B」, 「由 A 生 B」. 전이·전변: 「傳之 …」, 「變爲 …」. 시간적 진행: 「先 A 後 B」, 「秋傷於 A 冬生 B」. 사례(해수문): 「寒燥濕風火皆能令人咳」(동의 보감, COU-M09), 「五臟六腑皆令人咳 非獨肺也」(동의보감, COU-M12).

L3 치료 실행	구체적인 처치·개입의 방법·시점·순서·강 도·금기를 제시하는 층위. 지시(prescription)· 금기(prohibition)· 타이밍 조율의 서술이 일차 기능.	처치의 적정 지시, 금기·경고의 명시, 발병 단계별 분기, 처치의 순서 지정, 예후·난치 경고	처치 지시: 「宜 A」, 「以 A 治之」, 「用 A」, 「當 A」. 금기: 「忌 A」, 「禁 A」, 「勿 A」, 「不可 A」. 시점 분기: 「未發 … 已發 …」, 「初起 … 久病 …」, 「急用 … 緩用 …」. 순서: 「先 A 然後 B」. 예후 경고: 「A 者難治」, 「A 者死」. 사례(해수문): 「凡喘未發以扶正氣爲主, 已發以散邪爲主」(동의보감, COU-M40), 「凡火嗽忌用人蔘半夏陳皮等燥藥」(동의보감 , COU-M45).
----------------	---	--	--

판별 기준의 적용 원칙은 다음과 같다. 첫째, 층위 배정은 한 데이터 행의 일차적 기능(primary function)에 따른다. 동일한 한문 구절 안에서 정의·감별·처치의 어구가 혼재할 경우, 본 프로젝트의 1행 1판단 단위 원칙(3.6.1절)에 따라 별도 행으로 분리되며, 각 행은 그 일차적 기능에 따라 층위가 배정된다. 둘째, 어구 패턴의 출현이 층위 배정의 충분조건은 아니다. 예컨대 ‘忌’ 자가 등장한다고 해서 자동으로 L3에 배정되는 것이 아니라, 해당 구절의 기능적 역할이 처치 지시인지 감별 논리인지를 함께 검토한다. 가령 경계·정충·건망문의 ‘若作痰治 往往殺人(만약 담으로 보고 치료하면 종종 사람을 죽인다)’은 처치 금기의 형식을 띠지만, 해당 구절의 일차적 기능은 내인·외인 감별 판단을 명시하는 데 있으므로 L2(판단 구조)에 배정된다. 셋째, 문헌별 서술 양식의 차이로 인한 층위 교차 사례는 데이터셋의 Layer 필드에 그대로 기록되어 분석이 가능하도록 설계되었다.

이 판별 기준은 연구자의 해석적 판단을 기계적으로 대체하기 위한 것이 아니다. 한문 텍스트의 의미론적 다층성을 고려할 때 완전히 자동화된 층위 배정은 불가능하며, 본 표는 연구자의 텍스트 독해를 정밀화하고, 다른 연구자가 동일한 텍스트를 분석할 때 참조할 수 있는 명시적 근거를 제공하기 위한 운용적 기준(operational criteria)으로 작성되었다. 이러한 명시적 기준의 제시는 후술하는 5.3절의 다해석 병렬 수용 스키마(Interpreter_ID·Interpretation_Variant 필드)의 도입과 함께, 본 데이터 모델의 재현 가능성을 강화하는 두 축을 이룬다.

표 3. L1·L2·L3 판단 층위의 정의

층위	명칭	판단의 핵심 질문	텍스트 내 대응 요소
L1	현상 인식	이것을 어떤 질병 범주로 인식할 것인가	병증의 정의, 증상 기술, 계절·체질·환경 맥락

L2	판단 구조	어떤 원인이 관여하며 어떻게 감별할 것인가	병인론, 원인의 배열, 감별 기준, 병의 전변 경로
L3	치료 실행	언제, 어떻게 개입할 것인가	처방 선택, 개입 시점, 금기, 절제와 멈춤의 판단

이 세 층위가 반드시 텍스트 내에서 순차적으로 등장하는 것은 아니다. 『동의보감』은 대체로 L1→L2→L3의 순서를 따르는 경향이 있으나, 『의방유취』는 복수의 의서를 병렬적으로 인용하기 때문에 동일 병문 안에서 L1·L2·L3이 교차하여 나타나기도 한다. 『인제지』는 양생과 금기를 별도로 편성하여 L3의 범위가 치료 이후의 영역까지 확장된다. 이러한 문헌별 차이는 데이터셋의 Layer 필드에 기록되어 분석이 가능하도록 설계하였다.

해수문(咳嗽門)의 실제 데이터를 통해 L1→L2→L3 배정의 구체적 과정을 예시한다. 해수문은 RDM_Core 56행으로 구성되며, 『인제지』와 『동의보감』의 원문에서 추출한 대표 문장들은 다음과 같이 층위가 배정되었다.

(1) L1(현상 인식)의 배정 사례

『인제지』 해수문 서두에는 다음의 서술이 나온다. “咳謂無痰而有聲 肺氣傷而不清也(해는 담이 없고 소리만 있는 것으로 폐기가 상하여 맑지 않은 것이다)”, “嗽謂無聲而有痰 脾濕動而爲痰也(수는 소리 없이 담만 있는 것으로 비습이 동하여 담이 된 것이다)”, “咳嗽者有痰而有聲 因傷肺氣動於脾濕 故咳而兼嗽也(해수는 담과 소리가 함께 있는 것으로 폐기 손상과 비습 동요가 합작한 것이다)”. 이 구간은 해수라는 병증의 범위와 정의를 규정하고 있으므로 L1(현상 인식)에 배정된다. 데이터셋에서 각각 COU-M01, COU-M02, COU-M03으로 코딩되며, Provenance_Citation은 ‘河間方’이다.

(2) L2(판단 구조)의 배정 사례

『동의보감』 해수문 병인 부분에는 “秋傷於濕 冬生咳嗽(가을에 습에 상하면 겨울에 해수가 생긴다, COU-M06)”, “寒燥濕風火皆能令人咳(한·조·습·풍·화 모두 기침을 일으킬 수 있다, COU-M09)”, “五臟六腑皆令人咳 非獨肺也(기침은 폐만이 아니라 오장육부 모두에서 생긴다, COU-M12)”가 이어진다. 이 구간은 해수의 원인을 복수로 배열하고 장부 간 감별·전이의 기준을 제시하고 있으므로 L2(판단 구조)에 배정된다.

(3) L3(치료 실행)의 배정 사례

『동의보감』 해수문 치법 부분에는 “凡喘未發 以扶正氣爲主 已發 以散邪爲主(미발 시 정기를 돕고, 이미 발하면 사기를 흩는다, COU-M40)”, “凡火嗽 忌用人蔘半夏陳皮等燥藥

(무릇 화수에는 인삼·반하·진피 같은 조열한 약을 금한다, COU-M45)”이 나온다. 이 구간은 개입 시점에 따른 처치 분기와 구체적 금기를 명시하고 있으므로 L3(치료 실행)에 배정된다.

표 4. L1·L2·L3 배정 예시

현상코드	원문(발췌)	번역문(요지)	층위	배정 근거
COU-M03	咳嗽者有痰而有聲 因傷肺氣動於脾濕	폐기손상+ 비습 합작	L1	병증의 정의 규정
COU-M06	秋傷於濕 冬生咳嗽	가을 습→겨울 해수	L2	계절·환경 병인 서술
COU-M09	寒燥濕風火皆能令人咳	다원적 원인 배열	L2	복수 원인의 체계적 배열
COU-M12	五臟六腑皆令人咳 非獨肺也	오장육부 모두 관여	L2	장부 분기 판단 기준
COU-M40	凡喘未發 以扶正氣爲主 已發 以散邪爲主	미발=부정 / 이발=산사	L3	개입 시점별 처치 분기
COU-M45	凡火嗽 忌用人蔘半夏陳皮等燥藥	화수에 조약 금지	L3	금기의 명시적 서술

3.4 병문 유형 분류 체계(Type A-D)의 귀납적 도출

30병문을 L1·L2·L3 구조로 분절하고 병문 간 관계 데이터를 구축하는 과정에서, 각 병문이 조선 의학 판단 체계 안에서 수행하는 기능적 역할에 따라 네 가지 유형이 귀납적으로 드러났다. 이하에서는 각 유형의 조작적 정의, 분류 기준, 해당 병문을 제시한다.

Type A (생태순환형)

○정의: 인체와 외부 환경(계절·기후·지형)의 순환적 상호작용이 판단의 주된 축을 이루는 병문. L2 층위에서 병인이 ‘몸 내부의 기전’보다 ‘몸-환경 사이의 관계’로 서술되는 비율이 높은 것이 특징이다.

○분류 기준: 해당 병문의 L2 데이터에서 (a) 계절·기후·지형 등 환경 요인이 병인의 일차적 맥락으로 제시되고, (b) 치료 판단(L3)에서도 환경 변화에 따른 처방 조절이 핵심 논점을 이루는 경우.

○해당 병문(7개): 습병·조열·부종·해수·중한·고랭·중서.

→ 습병문 데이터에서 “風雨襲虛 山澤蒸氣 冒雨行濕 久臥濕地(DMP-SRC-03)”, 또는 해수문 데이터에서 “秋傷於濕 冬生咳嗽(COU-M06)”의 경우

Type B (판단노드형)

- 정의: 즉각적 개입의 여부와 방향을 결정하는 급성 분기점의 성격을 가진 병문. L2에서 감별 판단이 짧고 명확하며, L3에서의 개입이 신속한 결정을 요구하는 것이 특징이다.
- 분류 기준: 해당 병문의 L2 데이터에서 (a) 감별의 분기가 비교적 명확하고 (급성 vs. 만성, 독 vs. 비독 등), (b) L3에서 개입의 지연이 위험을 초래하는 것으로 서술되는 경우.
- 해당 병문(6개): 해독·금창·제상·곽란·중서·온역.
→ 곽란문 데이터에서 “時間性은 周期가 아니라 速度(GWR-SYS-01)” 또는 “急用吐法(GWR-SYS-03, 건곽란 응급)”의 경우

Type C (네트워크허브형)

- 정의: 다른 병문으로의 전이(轉移) 경로를 내포하는 복합 병리의 중심축 역할을 하는 병문. L2에서 단일 병인으로 수렴하지 않고 복수의 전변 경로가 서술되며, 데이터셋에서 NETWORK 시트의 ‘transitsTo’, ‘evolvesTo’ 관계가 다수 존재하는 것이 특징이다.
- 분류 기준: 해당 병문의 (a) L2 데이터에서 3개 이상의 다른 병문으로의 전이·전변이 서술되고, (b) NETWORK 관계 데이터에서 타 병문과의 연결(‘TRANSITS_TO’, ‘EVOLVES_TO’, ‘SHARES_PATHOLOGY’)이 25개 이상인 경우.
- 해당 병문(8개): 담음·혈병·제기·적취·허로·창만·내상·구토.
→ 허로문의 경우 “虛損→五勞→六極→七傷”이 전변 과정에서 4단계 심화 구조(허-07, 허-15, 허-21, 허-28)를 보이고, NETWORK 엣지 수는 46건이다.

Type D (경계판단형)

- 정의: 몸의 안팎을 연결하는 기관(눈·귀·코·입·소변·대변 등)에서 이루어지는 판단을 다루는 병문. L1에서 병증의 인식이 특정 기관의 생리적 경계 지점에서 출발하며, L2에서 내인과 외인의 감별이 핵심 판단을 이루는 것이 특징이다.
- 분류 기준: 해당 병문의 (a) L1 데이터에서 병증 인식이 특정 감각기관 또는 배설기관의 경계 기능과 결부되어 서술되고, (b) L2에서 동일 증상의 내인·외인 감별이 판단의 중심축을 이루는 경우.
- 해당 병문(9개): 안·이·비·치·소변·대변·황달·소갈·전간전광·경계정충건망.
→ 안문의 “眼爲臟腑之精(ANM-SRC-01)”의 경우 눈의 경계 기능이 장부와 관계 있음을 명시하고 있다.

경계 사례의 처리

위의 네 유형은 30병문 대부분에 대해 비교적 명확하게 적용되지만, 일부 병문에서는 복수의 유형에 걸치는 경계적 특성이 관찰되어 분류 판단이 단순하지 않았다. 분류 체계의 투명성을 위해 대표적인 경계 사례 세 건과 그 판단 근거를 밝힌다.

① 해수문(咳嗽門): Type A(생태순환형) vs. Type C(네트워크허브형).

해수문은 Type A로 분류하였으나 Type C의 특성도 뚜렷하게 공존한다. NETWORK 시트에는 장부 간 전이 관계가 다수 기록되어 있다. 폐해→대장해(COU-M13, M20), 비해→위해(COU-M16, M18), 신해→방광해(COU-M17, M21), 久咳→삼초해(COU-M22) 등 TRANSITS_TO 엣지가 4건이며, 답음문과의 접속 경로(脾積濕→痰→上入於肺→嗽)도 LEADS_TO 체인으로 명시되어 있다. 총 NETWORK 엣지 수도 27건으로 Type C 병문의 평균(25건)에 근접한다. 그러나 최종적으로 Type A로 분류한 근거는 다음과 같다. 해수문의 L2 데이터에서 판단의 출발점이 되는 서술들은 “秋傷於濕 冬生咳嗽(COU-M06)”, “寒燥濕風火皆能令人咳(COU-M09)” 등 계절·기후·환경 요인이 병인의 일차적 맥락으로 제시되는 구간이다. 장부 간 전이는 이 환경-인체 상호작용이 오래 지속될 때(“久咳不已”) 발생하는 이차적 경로로 서술되고 있다. 즉 해수문의 일차적 판단 기능은 ‘몸-환경 순환 관계의 인식’에 있으며, 타 병문으로의 전이는 이 일차 판단의 후속 전개에 해당한다. 반면 답음문은 처음부터 복수 병문(해수·구토·현훈·심계·전간전광)으로의 전이를 핵심 구조로 삼고 있어, 동일한 전이 현상이라도 병문 안에서의 구조적 위상이 다르다.

② 소갈문(消渴門): Type D(경계판단형) vs. Type C(네트워크허브형).

소갈문은 Type D로 분류하였으나, NETWORK 엣지 수가 29건으로 30병문 중 4위에 해당하며, 내부에 상소→중소→하소→강증으로 이어지는 단계적 심화 구조(TRANSITIONS_TO 체인)와 소갈→뇌저/배창, 소갈→중만/고창, 소갈→옹저 등 타 병증으로의 전변 경로가 다수 기록되어 있어 Type C의 특성이 강하게 나타난다. 그러나 최종적으로 Type D로 분류한 근거는 다음과 같다. 소갈문의 L1 데이터는 ‘갈증(渴)·다음(多飲)·다뇨(多尿)·다식(多食)’이라는 신체 경계 기관(입, 위장, 방광)의 출입 이상을 병증 인식의 출발점으로 삼고 있다. L2의 핵심 구조인 삼소 분기(上消=폐, 中消=위, 下消=신) 역시 각 장부의 ‘경계 기능’(진액 출납, 수곡 운화, 정수 배설)이 판단의 축을 이룬다. 타 병증으로의 전변은 소갈이 오래 경과한 뒤의 예후적 전개(SGM-46~48)로서, 병문의 일차적 판단 기능 자체는 몸의 안팎을 연결하는 경계 지점에서 이루어지는 감별에 있다. 한편 소갈문의 높은 NETWORK 엣지 수는, 경계판단형 병문이 반드시 단순한 구조를 갖는 것은 아니며 경계 기관을 매개로 복수의 장부와 연결될 수 있음을 보여주는 사례이기도 하다.

③ 경계·정충·건망문(驚悸·怔忡·健忘): Type D(경계판단형) vs. Type C(네트워크허브형).

이 병문은 Type D로 분류하였으나, NETWORK 시트에 驚悸→怔忡(EVOLVES_TO), 怔忡→健忘(EVOLVES_TO)이라는 명확한 시간적 전환 구조가 기록되어 있으며, 五飲停蓄→痰飲病門(LINKS_TO), 肝腎虛/心氣弱→虛勞病門(LINKS_TO)과 같은 타 병문으로의 접속 경로도 존재한다. 이러한 진행형 구조와 허브적 연결은 Type C의 특성에 가깝다. 그러나 최종적으로 Type D로 분류한 근거는 다음과 같다. 경계·정충·건망의 세 병증은 모두 ‘심(心)’이라는 장부의 경계 기능(神明(神明)의 안정, 정서적 자극에 대한 반응, 인지의 유지)에서 출발하는 판단이다. L2의 핵심 분기축은 심허+담울(→경계), 심허+정수(→정충), 심비혈소+신후(→건망)로서, 모두 내인(심허, 혈소)과 외인(외부 충격, 담울)의 감별이 판단의 중심을 이룬다. 驚悸→怔忡→健忘의 시간적 전환은 하나의 경계판단이 만성화되는 과정을 보여주는 것이지, Type C처럼 질적으로 다른 병문으로 전이되는 구조는 아니다. 또한 이 병문의 핵심 경고인 “若作痰治 往往殺人(만약 담으로 치료하면 종종 죽인다, 인제지)”은 경계 기관(심)의 내인적 허손을 외인적 담병으로 오인하는 것을 금하는 서술로서, 내인·외인 감별이 이 병문의 일차적 판단 기능임을 확인시켜 준다.

이상의 세 사례에서 확인되듯, Type A-D 분류에서 가장 빈번하게 발생하는 경계 판단은 ‘Type C(네트워크허브형)와의 중첩’이다. 이는 조선 의학의 병문들이 고립적으로 존재하지 않고 상호 전이의 네트워크 안에 놓여 있기 때문이다. 본 연구의 분류 기준은 이러한 네트워크적 속성의 유무가 아니라, 해당 병문의 일차적 판단 기능이 어디에 놓이는가에 따라 유형을 결정하였다. 전이 경로의 존재 자체가 아니라, 전이가 병문의 중심 구조를 이루는가 아니면 일차 판단의 후속 전개에 해당하는가를 기준으로 삼은 것이다.

3.5 Provenance_Citation 설계와 문헌 계보 보존

3대 원칙 중 분리 원칙의 기술적 실현이 Provenance_Citation 설계다. 조선 의학 문헌은 자체적인 서술과 함께 중국 의서를 광범위하게 인용하고 있다. 예를 들어 『동의보감』의 해수문(咳嗽門)이 『내경(內經)』이나 『하간(河間)』의 특정 구절을 인용하는 경우, 그 인용의 출처가 데이터 안에 기록되어야 한다.

Provenance_Citation 필드는 각 데이터 행이 어떤 문헌의 어떤 구절에 근거하는지를 기록한다. 이 설계는 두 가지 기능을 수행한다. 첫째, 동일한 병증에 대해 서로 다른 문헌이 제시하는 판단을 비교할 수 있게 한다. 둘째, 조선 문헌이 인용한 중국 의서의 계보까지 추적함으로써 동아시아 의학 지식의 전송 경로를 데이터 안에 보존한다. 이 설계는 오류와 진본을 분별하는 데에도 기여한다. 동일한 원문이 『의방유취』와 『동의보

감』에서 서로 다르게 인용된 경우, Provenance_Citation의 교차 대조를 통해 이를 식별할 수 있다.

3.6 데이터 스키마의 뼈대

각 병문의 RDM 데이터셋은 복수의 시트로 구성된 하나의 패키지(AllInOne) 형태를 취한다. 30병문 전체에 걸쳐 공통적으로 사용되는 시트 구조는 다음과 같다.

표 5. RDM 데이터셋의 시트 구성

시트	기능	성격
RDM_Core	병문 내부의 판단 단위를 행 단위로 기록	핵심 데이터
RDM_Canonical	Core 데이터를 모문장 단위로 정렬·정제한 뷰	정제 뷰
RDM_Parent	판단 단위를 상위 모듈로 묶는 구조	상위 구조
FORMULA_LAYER	처방·약재조합·적응상황을 기록	치료 데이터
STAGE3_BRIDGE	『역시만필』 등 임상 사례와의 접속 경로 기록	임상 연결
NETWORK	병문간·병문내의 관계를 노드·엣지로 기록	관계 데이터
SCHEMA	해당 병문 데이터셋의 컬럼 정의와 입력 원칙	메타데이터
CODEBOOK/CODE_GUIDE	현상코드·대분류·Layer 등의 코드 정의표	메타데이터
FIGURE	병문 사고 구조의 텍스트 도식	구조 도식
README	버전 이력·사용 문헌·보장 사항 기록	문서 정보

이 중 가장 핵심적인 시트는 RDM_Core와 NETWORK다. RDM_Core는 병문 내부의 판단 구조를 기록하고, NETWORK는 병문 간(및 병문 내부)의 관계를 기록한다.

3.6.1 RDM_Core 시트의 컬럼 구조

30병문의 데이터 구축은 약 1년에 걸쳐 진행되었으며, 이 과정에서 스키마가 점진적으로 발전하였다. 현재 데이터셋에는 두 가지 스키마 패턴이 공존한다.

표 6. RDM_Core 스키마의 두 가지 패턴

구분	초기형(14 컬럼)	정제형(11~15 컬럼)
----	------------	---------------

적용 병문	허로, 제기, 혈병, 경계정충견망 등	해수, 소갈, 황달, 담음 등
판단 층위 표기	대분류 필드 (병원/형증/맥법/오로 등)	Layer 필드 (L1/L2/L3)
번역문 필드명	해석	번역문
문헌 위치	문헌 3 축 위치+ 조선문헌 좌표 (2 필드)	문헌축 + 문헌 출전 (2 필드)
병리 정보	주 장부통로, 병리기전, 시간성 (3 필드)	구조축 (1 필드에 통합)
개입 정보	개입위치 (별도 필드)	전환 필드에 통합
확장 필드	-	모문장 ID, 판단군, Source_Citation, 연결병문 (Canonical 시트)

두 패턴의 차이는 표현 방식의 차이이지 기록하는 정보의 본질적 차이는 아니다. 양자 모두에서 공통적으로 유지되는 핵심 속성은 하단의 7가지다.

① 판단 층위(Layer/대분류)

각 데이터 행이 L1(현상 인식)·L2(판단 구조)·L3(치료 실행) 중 어디에 해당하는지를 표시한다. 초기에는 ‘병원’, ‘형증’, ‘맥법’, ‘오로’, ‘치법’ 등 병문 고유의 대분류명으로 기록하였으며, 이들은 L1·L2·L3과 대응 관계에 있다(예: 형증·맥법→L1, 병원·오로·감별→L2, 치법·분류치료→L3).

② 현상코드

각 판단 단위의 고유 식별 코드다. 병문별 접두어와 일련번호로 구성된다(예: COU-M01, 허-15, SGM-03, ANM-SRC-08). 동일 병문 내에서 행을 고유하게 식별하고, NETWORK 시트에서 근거행(Evidence)으로 참조된다.

③ 원문(한문)

해당 판단 단위의 근거가 되는 한문 원전 텍스트다. 분리 원칙에 따라 번역문과 별도 필드로 기록된다. 1행 1판단 단위를 원칙으로 하되, 텍스트의 성격에 따라 복수 문장이 하나의 행에 포함되기도 한다.

④ 번역문(한글)

원문에 대응하는 한글 번역이다. 기존 역주본과 〈한의학고전DB〉의 번역을 참조하되 최종적으로 연구자의 판단에 의해 작성되었다.¹⁸

⑤ 문헌축/문헌 출전

해당 데이터 행이 3축(『의방유취』·『동의보감』·『인제지』) 중 어느 문헌에서 추출되었는지, 그리고 해당 문헌 내부의 어느 편·문·조항에 위치하는지를 기록한다.

⑥ Provenance_Citation

조선 편찬서가 명시한 상위 인용 출처다. 예컨대 『동의보감』 해수문이 ‘河間方’이나 ‘內經’을 인용한 경우 그 출처명이 기록된다. 이 필드는 ‘문헌축’과 구별된다. “이 데이터가 어느 조선 문헌에서 왔는가”를 기록하는 것이 문헌축이라면, Provenance_Citation은 “그 조선 문헌이 어느 중국 의서를 인용하고 있는가”를 기록하는 것이다. 3.5절에서 논의한 문헌 계보 보존의 기술적 실현이 곧 이 필드라 하겠다.

⑦ 전환/비고

해당 판단 단위에서 다른 판단 단위 또는 다른 병문으로의 전이·전변 경로, 금기 사항, 주의 사항 등을 기록한다. 화살표 표기(예: “肺氣傷+脾濕動→咳嗽”, “허손→허로”)로 간결하게 적는 것을 원칙으로 한다.

Layer	현상코드	대묘물	구조축	원문	번역문	문헌축	문헌출전	Provenance_Citation	금기
L1	COU-M01	해수(咳嗽) — 상태 순환형 병문	정의·구분	咳嗽無痰而有聲, 肺氣傷而不滿也	해는 달이 없고 소리만 있는 것으로, 폐기가 상하여 맑지 않은 것이다.	인제지	해수 형증·이름과 증상	河間方	肺氣傷→咳
L1	COU-M02	해수(咳嗽) — 상태 순환형 병문	정의·구분	嗽謂無聲而有痰, 脾濕動而為痰也	수는 소리 없이 다만 있는 것으로, 비습이 통하여 담이 된 것이다.	인제지	해수 형증·이름과 증상	河間方	脾濕動→痰
L1	COU-M03	해수(咳嗽) — 상태 순환형 병문	정의·구분	咳嗽者有痰而有聲, 因傷脾氣動於脾濕, 故咳而兼嗽也	해수는 담이 있어서 소리도 있는 것으로, 폐기 손상과 비습 증요가 함께 작용한 것이다.	인제지	해수 형증·이름과 증상	河間方	肺氣傷+脾濕→咳
L1	COU-M04	해수(咳嗽) — 상태 순환형 병문	외감개시	肺主氣, 形寒飲冷則傷之, 使氣上而不下, 逆而不收	폐는 기를 주관하나 형이 차고 냉음을 마시면 상하여 기가 위로 올라가고 내려오지 않는다.	인제지	해수 형증·이름과 증상	明理論	形寒飲冷→
L1	COU-M05	해수(咳嗽) — 상태 순환형 병문	방현형상	喉中漉漉如蟬, 胃弱如梗, 是冷嗽	목 안이 침침 간지럽고 갈갈하여 가시가 걸린 듯 하면 냉수이다.	인제지	해수 형증·이름과 증상	明理論	咽喉梗塞→
L2	COU-M06	해수(咳嗽) — 상태 순환형 병문	형성조건	秋傷於風, 冬生咳嗽	가을에 습해 상하면 겨울에 해수가 생긴다.	동의보감	咳嗽病因	內經	秋風→冬咳
L2	COU-M07	해수(咳嗽) — 상태 순환형 병문	형성기전	秋傷於風, 濕於脾也	가을의 습사는 비에 형여 겨울 기침의 기저가 된다.	동의보감	咳嗽病因	河間	秋風→脾積
L2	COU-M08	해수(咳嗽) — 상태 순환형 병문	병리기전	脾無留濕, 則雖傷脾氣, 亦不為痰	비에 습이 머물지 않으면 폐기를 상해도 담은 되지 않는다.	동의보감	咳嗽病因	河間	脾濕留滯→
L2	COU-M09	해수(咳嗽) — 상태 순환형 병문	원인분류	寒燥風火皆能令人咳嗽	한·조·습·풍·화는 모두 기침을 일으킬 수 있다.	동의보감	咳嗽病因	河間	六氣/病性
L2	COU-M10	해수(咳嗽) — 상태 순환형 병문	전이기전	濕病痰飲入胃, 留之而不行, 上入於肺則為嗽	습병의 담음이 위에 머물러 돌지 않다가 위로 폐에 오르면 수가 된다.	동의보감	咳嗽病因	河間	胃留痰飲→
L2	COU-M11	해수(咳嗽) — 상태 순환형 병문	분기판단	濕在心經, 謂之熱痰; 濕在肝經, 謂之風痰; 濕在脾經, 謂之氣痰; 濕在腎經, 謂之寒痰	습이 머무는 경맥에 따라 열담·풍담·기담·전담으로 일컫는다.	동의보감	咳嗽病因	河間	濕停經絡→
L2	COU-M12	해수(咳嗽) — 상태 순환형 병문	장부분기	五臟六腑皆令人咳嗽, 非獨肺也	기침은 폐만이 아니라 오장육부 모두에서 생긴다.	인제지	해수 장부전변	內經	多臟腑→咳
L2	COU-M13	해수(咳嗽) — 상태 순환형 병문	오장형증	肺咳之狀, 咳而喘急有音, 甚則唾血	폐는 기침하며 숨결에 소리가 나고 심하면 타열한다.	인제지	해수 장부전변	內經	肺受邪→咳
L2	COU-M14	해수(咳嗽) — 상태 순환형 병문	오장형증	心咳之狀, 咳則心痛, 喉中啞啞如蟬狀	심해는 기침하면 심통이 있고 목에 가시가 걸린 듯하다.	인제지	해수 장부전변	內經	心受邪→心
L2	COU-M15	해수(咳嗽) — 상태 순환형 병문	오장형증	肝咳之狀, 咳則兩脅下痛, 甚則不可以轉	간해는 기침하면 양脇 아래가 아프고 심하면 몸을 돌릴 수 없다.	인제지	해수 장부전변	內經	肝受邪→肝
L2	COU-M16	해수(咳嗽) — 상태 순환형 병문	오장형증	脾咳之狀, 咳則右脇下痛, 陰陰引肩背, 動則咳劇	비해는 우脇 아래가 아프고 어깨와 등이 당기며 움직이면 더 심해진다.	인제지	해수 장부전변	內經	脾受邪→右
L2	COU-M17	해수(咳嗽) — 상태 순환형 병문	오장형증	腎咳之狀, 咳則腰背相引而痛, 甚則遺尿	신해는 요배가 서로 당기며 아프고 심하면 담뇨를 뱉는다.	인제지	해수 장부전변	內經	腎受邪→腰
L2	COU-M18	해수(咳嗽) — 상태 순환형 병문	육부전이	脾咳不已, 則胃受之, 胃咳之狀, 咳而嘔	비해가 오래가면 담으로 옮겨 기침하며 구역한다.	인제지	해수 장부전변	內經	脾咳久→胃
L2	COU-M19	해수(咳嗽) — 상태 순환형 병문	육부전이	肝咳不已, 則膽受之, 膽咳之狀, 咳而膽汁	간해가 오래가면 담으로 옮겨 담즙을 토한다.	인제지	해수 장부전변	內經	肝咳久→膽
L2	COU-M20	해수(咳嗽) — 상태 순환형 병문	육부전이	肺咳不已, 則大腸受之, 大腸咳狀, 咳而遺矢	폐해가 오래가면 대장으로 옮겨 기침하며 대변을 지린다.	인제지	해수 장부전변	內經	肺咳久→大
L2	COU-M21	해수(咳嗽) — 상태 순환형 병문	육부전이	腎咳不已, 則膀胱受之, 膀胱咳狀, 咳而遺尿	신해가 오래가면 방광으로 옮겨 기침하며 유뇨한다.	인제지	해수 장부전변	內經	腎咳久→膀胱

그림 1. 실제 RDM_Core 시트의 예시

3.6.2 NETWORK 시트의 컬럼 구조

NETWORK 시트는 병문 간 및 병문 내부의 관계를 노드-엣지 형식으로 기록한다. 기

본 구조는 다음 4~6개 컬럼으로 구성된다.

표 7. NETWORK 시트의 컬럼 구조

컬럼명	설명	예시
Source (또는 From_Node, 노드 A)	관계의 출발 노드	咳嗽, 虛損, 驚悸
Edge (또는 Relation, Edge_Type)	관계 유형	TRANSITIONS_TO, CAUSES, SHARES_PATHOLOGY
Target (또는 To_Node, 노드 B)	관계의 도착 노드	痰, 虛勞, 怔忡
Evidence (또는 근거행)	해당 관계의 근거가 되는 RDM_Core 현상코드	COU-M06, 허-07, GB-002
NodeType (일부 병문)	노드의 유형 분류	Disease, Pathway, Diagnostic, Execution
설명 (일부 병문)	관계에 대한 부연 설명	"허손이 점차 심해지면 허로가 됨"

NETWORK 시트에서 사용된 관계 유형은 TRANSITIONS_TO, CAUSES, PART_OF, SHARES_PATHOLOGY, HAS_FORMULA, SPLITS_TO, BRIDGES_TO 등 37종에 달한다. 이 관계 유형들은 사전에 일괄 설정된 것이 아니라 개별 병문의 데이터 구축 과정에서 텍스트의 서술에 따라 귀납적으로 도출·추가되었다. Evidence 컬럼에 RDM_Core의 현상코드가 기록됨으로써, 각 관계가 어떤 원문 근거에 기반하는지를 역추적할 수 있다.

4. 편찬 데이터와 그 특징

이 장에서는 30병문 RDM의 구축 결과를 제시하고, 유형별 분포, L2 층위의 구조적 차이, 병문 간 관계의 패턴을 살펴본다. 이어서 데이터 예시를 통해 귀납적으로 확인되는 조선 의학의 두 구조적 특성에 관해 정리한다.

4.1 30 병문 데이터셋 개요와 DataON 등록 현황

2025년 2월부터 2026년 3월까지 구축된 30병문 데이터셋은 KISTI DataON 국가연구 데이터 플랫폼에 공식 등록되었다.¹⁹ 전체 데이터셋의 규모는 다음과 같다.

표 8. 30 병문 데이터셋의 전체 규모

항목	수치
----	----

분석 대상 병문 수	30
RDM_Core 총 데이터 행 수	약 750 행 이상
NETWORK 총 관계(엣지) 수	약 485 건 이상
L1(현상 인식) 비율	약 31%
L2(판단 구조) 비율	약 45%
L3(치료 실행) 비율	약 24%
병문당 평균 RDM_Core 행 수	약 25 행
병문당 평균 NETWORK 엣지 수	약 17 건
최대 RDM_Core 행 수	해수문 56 행
최대 NETWORK 엣지 수	허로문 46 건

L2(판단 구조) 비율이 약 45%로 가장 높다는 점이 주목할 만하다. 조선 의학 문헌이 증상 나열(L1)이나 처방 목록(L3)이 아니라, 판단의 논리를 서술하는 데 가장 많은 지면을 할애하고 있음을 보여주는 단서라 하겠다.

4.2 유형별 분포와 구조적 특성

30병문의 유형별 분포는 다음과 같다.

표 9. 30 병문의 유형별 분포

유형	병문 수	비율	평균 Core 행	평균 NET 엣지	대표병문
A 생태순환형	7	23.3%	28	16	해수, 습병, 중서
B 판단노드형	6	20.0%	26	14	곽란, 온역, 금창
C 네트워크허브형	8	26.7%	28	25	허로, 담음, 혈병
D 경계판단형	9	30.0%	27	15	안문, 소갈, 황달

Type C의 평균 NETWORK 엣지 수(25건)가 다른 유형(14~16건)에 비해 뚜렷하게 높다는 것을 알 수 있다. 유형 분류와 관계 데이터 사이의 정합성을 보여주는 근거라 하겠다.

4.3 병문 간 관계 데이터의 패턴 분석

NETWORK 시트에 기록된 약 485건의 관계 데이터에서 사용된 관계 유형은 CAUSES, TRANSITIONS_TO, PART_OF, HAS_FORMULA 등 30종 이상에 달한다. 이 유형들은 사전에 일괄 설정된 것이 아니라 개별 병문의 데이터 구축 과정에서 텍스트의 서술에 따라 귀납적으로 도출·추가되었기 때문에 명칭에 다소간의 변이가 있다. 분석의 편의를 위해 의미적 유사성에 따라 상위 범주로 묶어 정리하면 다음과 같다.

표 10. NETWORK 관계 유형 범주별 빈도

관계 범주	포함 유형(대표)	건수	비율
전이·분기	TRANSITIONS_TO, EVOLVES_TO, LEADS_TO, TRANSITS_TO, SPLITS_TO	약 90 건	약 19%
연결·교량	BRIDGES_TO, LINKS_TO, SHARES_PATHOLOGY, ASSOCIATED_WITH, COMBINES_WITH	약 90 건	약 19%
구성·포함	PART_OF, HAS_COMPONENT, HAS_STAGE, HAS_PHENOMENON, INCLUDES	약 80 건	약 16%
인과·생성	CAUSES, GENERATES, TRIGGERS	약 70 건	약 14%
처방·실행	HAS_FORMULA, TREAT_AS, REQUIRES, TREATED_BY, INTERVENES_IN	약 50 건	약 10%
진단·표지	INDICATES, HAS_MARKER, HAS_MECHANISM, SIGNAL	약 22 건	약 5%
감별·대비	DISTINGUISH_FROM, CONTRASTS_WITH, PARALLEL_WITH, HAS_DIAGNOSTIC_ETHIC	약 11 건	약 2%
금기·경고	AVOID, CONTRAINDICATES, FORBIDS, RISKS	약 10 건	약 2%
예후	HAS_OUTCOME_RULE, HAS_WARNING	약 5 건	약 1%
기타	(분류 미상)	약 57 건	약 12%
합계		약 485 건	100%

이 분포에서 주목할 점은 세 가지다.

① 가장 빈도가 높은 관계 범주는 ‘전이·분기’와 ‘연결·교량’이다. 이 두 범주가 전체의 약 38%를 차지한다. 이는 조선 의학이 개별 병문을 독립적으로 파악하지 않고, 전이 경로(A가 B로 전변한다), 병리적 교량(A와 B가 기전을 공유한다)의 맥락 속에서 질병을 이해했음을 보여주는 근거라 하겠다. 반면 ‘금기·경고’(약 2%)와 ‘예후’(약 1%)의 비율이 낮은 것은, 금기와 예후 판단이 NETWORK의 병문 간 관계보다는 RDM_Core 내부의 L3 데이터에 주로 기록되어 있기 때문이다. 한편 ‘구성·포함’(PART_OF 등)이 약 16%에 달하는 것은, NETWORK 시트가 병문 ‘간’의 관계만을 기록하는 것이 아니라 병문 ‘내부’의 하위 구조(예: 제기문 내부의 상기·기역·중기·하기·단기 등)를 노드화하여 기록하는 역할도 겸하고 있음을 반영한다. 이 설계 덕분에 병문 내부의 세부 구조도 그래프 탐색의 대상이 될 수 있다.

② 특정 병문이 NETWORK 관계의 허브 역할을 수행한다. NETWORK 엣지 수 상위 5개 병문은 다음과 같다.

표 11. NETWORK 엣지 수 상위 5개 병문

순위	병문	유형	NETWORK 엣지 수	타 병문 참조 수
----	----	----	--------------	-----------

1	허로문(虛勞)	C	46 건	-
2	제기문(諸氣)	C	30 건	6 건
3	혈병문(血病)	C	30 건	8 건
4	소갈문(消渴)	D	29 건	-
5	해수문(咳嗽)	A	27 건	-

허로문이 46건으로 압도적이다. 허로문의 NETWORK에는 단계적 전이 (TRANSITIONS_TO: 虛損→虛勞, 五勞→六極, 六極→七傷), 인과 관계(CAUSES: 寒感虛損→陽虛, 熱感虛損→陰虛, 房勞→陰血虛), 감별 관계(DISTINGUISH_FROM: 氣虛↔血虛, 陽虛↔陰虛), 처방 연결(HAS_FORMULA: 陰虛→大補陰丸, 陽虛→桂附湯, 陰陽俱虛→十全大補湯), 예후 경고(HAS_WARNING: 不受補者難治, 五敗死證) 등 거의 모든 관계 범주가 골고루 포함되어 있어, 허로문이 단순히 연결이 많은 것이 아니라 관계의 다양성 면에서도 가장 풍부한 허브임을 보여준다. 제기문 역시 SHARES_PATHOLOGY를 통해 혈병문(기불섭혈, 기체→혈어), 조열문(기울화화), 해학문(외감기체), 중한문(기폐) 등 4개 이상의 타 병문과 병리 기전을 공유하는 것으로 기록되어 있다.

③ Type C(네트워크허브형)로 분류된 병문이 실제로 높은 연결 수를 보인다. 유형별 평균 NETWORK 엣지 수를 비교하면 다음과 같다.

표 12. 유형별 평균 NETWORK 엣지 수

유형	평균 NETWORK 엣지 수	합계
Type A 생태순환형	16.2 건	97 건
Type B 판단노드형	11.2 건	67 건
Type C 네트워크허브형	22.1 건	177 건
Type D 경계판단형	14.4 건	144 건

Type C의 평균(22.1건)이 다른 유형(11.2~16.2건)에 비해 뚜렷하게 높으며, 엣지 수 상위 5개 병문 중 3개(허로문, 제기문, 혈병문)가 Type C에 해당한다. 이는 3장에서 귀납적으로 도출한 유형 분류가 실제 관계 데이터와 정합적임을 사후적으로 검증하는 결과다. Type B(판단노드형)의 평균이 가장 낮은 것(11.2건) 역시 이 유형의 특성(급성 분기점으로서 신속한 판단을 요구하며, 복잡한 전이 네트워크보다는 명확한 이항 분기를 특징으로 함)과 부합한다.

한편 Type D인 소갈문이 29건으로 상위에 위치한다는 사실은 주목할 만하다. 이는 3.4절의 경계 사례 분석에서 논의한 바와 같이, 소갈문이 경계판단형이면서도 상소→중소→하소→강증의 단계적 전이 구조와 용저·뇌저·고창 등 타 병증으로의 전변 경로를 풍

부하게 갖추고 있기 때문이다. 유형 분류는 병문의 ‘일차적 판단 기능’에 따른 것이므로, 높은 연결 수가 반드시 Type C 분류를 의미하지는 않는다.

이상의 분석을 종합하면, 30병문의 NETWORK 데이터는 세 가지를 보여준다. 첫째, 조선 의학은 병문을 독립적으로 다루지 않고 인과·전이·공유의 관계망 안에서 파악한다. 둘째, 이 관계망에는 허로문·제기문·혈병문과 같은 허브 병문이 존재하며, 이들이 복합 병리의 교차점 역할을 수행한다. 셋째, Type A-D 유형 분류와 관계 데이터 사이에 통계적 정합성이 확인된다.

4.4 데이터에서 귀납된 두 특성: 원인배열과 개입타이밍 조율

상기의 분석 결과를 종합하면, 30병문 데이터 전체에 걸쳐 두 가지 구조적 특성이 일관되게 확인된다.

4.4.1 원인배열의학: 병을 단일 원인으로 환원하지 않는다

30병문의 L2(판단 구조) 데이터를 통해 확인되는 첫 번째 특성은, 조선 의학이 병의 원인을 단일 요인으로 환원하지 않고 복수의 원인을 함께 체계적으로 배열한 위에서 판단을 내린다는 점이다. 해수문(咳嗽門)을 예로 들면, 조선 의학은 기침이라는 현상에서 출발하되 그것을 폐의 문제로만 단지 않는다. “寒燥濕風火皆能令人咳(한·조·습·풍·화 모두 기침을 일으킬 수 있다, COU-M09)”는 원인의 다원성을 선언하고, “五臟六腑皆令人咳 非獨肺也(기침은 폐만이 아니라 오장육부 모두에서 생긴다, COU-M12)”는 장부 귀속의 단일 환원을 명시적으로 거부한다. 이어서 담(痰)의 관여 여부, 외감과 내상의 감별, 허증과 실증의 판단, 시간적 전변의 방향이 함께 배열되고, 그 배열 위에서 판단이 구성된다. 이러한 특성은 해수문에 국한되지 않는다. 유형별 대표 병문의 L2 데이터에서 확인되는 원인배열 구조를 정리하면 다음과 같다.

표 13. L2 층위에서 확인되는 원인배열 구조의 사례

병문	유형	L2에서 배열되는 원인 축	대표 원문 근거	배열 축 수
해수문	A	외감/내상, 한/조/습/풍/화, 오장육부 분기, 시간적 전변	COU-M06, M09, M12, M18~22	4
습병문	A	외습/내습, 풍습/한습/습열, 계절·지형, 비허 기저	습병 RDM L2 데이터	4
곽란문	B	음식손상, 내적외감, 풍습갈 합병	GWR-SRC-02~04	3
허로문	C	한감손양/열감손음, 방로·사려/노역·기아, 상하 손상 방향, 장부별 분기(오로)	허-04~07, 허-15, 허-37	5

답음문	C	열→담/수→음, 비허→수액정체, 정맥별 분기(열담/풍담/기담/한담), 장부 귀류	DM-03~04, DM-11	4
제기문	C	방향 이상(상기/기역/중기/하기), 연속성 이상(단기/소기), 응결 이상(기체/기울/기적), 원인 엔진(칠기/구기/음식상/외감)	제기문 FIGURE 구조도	4
안문	D	육음외상, 오장내울, 음식·방로, 원시·비읍(悲泣), 초사(抄寫)·조루(雕鏤)	ANM-SRC-08	5
소갈문	D	위·대장 열결, 진액·혈 부족, 고량비감 식이, 삼소 장부 분기(폐/위/신)	SGM-01~10, SGM-15~17	4
경계·정충문	D	심허+ 담울, 심허+ 정수, 간신허+ 심기약, 심비혈소+ 신희	경계 NETWORK GB-004~008	4

9개 대표 병문 모두에서 배열 축 수가 3 이상이다. Type B(곽란문)에서조차 음식손상, 내적외감, 풍습갈 합병이라는 3개 축이 배열되며, Type C(허로문)와 Type D(안문)에서는 5개 축이 확인된다. 30병문 전체를 검토한 결과, 배열 축 수가 2 이하인 병문은 확인되지 않았다. 이는 원인배열이 특정 병문의 예외적 특징이 아니라 조선 의학 판단 체계 전체의 구조적 특성임을 보여준다.

이 특성을 본 연구는 ‘원인배열의학(Cause-Arrangement Medicine)’으로 명명한다. 원인배열의학이란 병의 원인을 단일 요인으로 축소(귀속, 환원, 단정)하지 않고, 관련된 복수의 원인을 체계적으로 배열하여 그 관계 속에서 판단을 구성하는 의학적 사유 방식을 뜻한다. 여기서 ‘배열’이라는 표현에 주목할 필요가 있다. 배열은 단순히 나열(listing)되는 것이 아니라, 일정한 순서와 관계 속에 감별과 판단 또는 방향을 전제로 배치된다는 뜻이다. 해수문에서는 외감과 내상이 먼저 감별되고, 그 위에서 허와 실이 판단되며, 시간적 전변의 방향이 고려된다. 허로문에서는 한감(寒感)과 열감(熱感)이 먼저 분별되고, 그에 따라 손상의 방향이 상하로 갈리며(허-04: “感寒則損陽 損自上而下”, 허-05: “感熱則損陰 損自下而上”), 이 방향 위에서 장부별 귀속이 결정된다(허-06). 이 순서 자체가 판단의 구조라 할 수 있다. 이것이 ‘원인나열’이 아니라 ‘원인배열’로 명명하는 이유다.

4.4.2 개입타이밍 조율의학: 언제 개입하고 언제 멈추는가

두 번째 특성은 개입의 타이밍에 극도로 민감하다는 점이다. 이 특성은 주로 L3(치료 실행) 층위에서 확인되나, L2(판단 구조)에서 병의 단계적 진행이 서술될 때 이미 그 전제가 마련된다. 허로문(虛勞門)은 이 특성이 가장 뚜렷하게 드러나는 병문이다. 허로는 허손이 누적되어 오장·기혈·음양·정기가 단계적으로 붕괴하는 진행형 구조를 가진다. 동일한 허로라 하더라도 1손(一損)의 단계와 5손(五損)의 단계에서 개입의 방법이 근본적으로 달라진다. 초기에는 적극적 보강이 가능하나, 말기에는 “不受補者難治(보약을 받아들이지 못하면 난치, 허-49)”라 하여 오히려 개입을 자제하고 경과를 살피는 것이 최선으로 서술된다. 나아가 “峻補燥熱(준보·조열 약 남용 금지, 허-40)”라 하여 성급한 개입

이 해악이 됨을 명시적으로 경고한다. 담음문(痰飲門) 역시 마찬가지다. 담음은 해수·구토·현훈·심계·전간전광·중풍 등으로 전이되는 체액 병리 네트워크의 허브 역할을 하며, 같은 담음이라도 어느 경로로 전이 중인가에 따라 처치가 달라진다. “先逐去敗痰 然後看虛實調理(먼저 패담을 몰아내고 그 뒤 허실을 살펴 조리한다, DM-08)”는 개입의 순서가 결과를 좌우함을 보여주는 서술이다. 이러한 타이밍 조율은 복수의 병문에 걸쳐 광범위하게 확인된다. 유형별 대표 병문의 L3 데이터에서 확인되는 사례를 정리하면 다음과 같다.

표 14. L3 층위에서 확인되는 개입타이밍 조율의 사례

병문	유형	타이밍 관련 판단의 핵심	대표 원문 근거	타이밍 조율 양상
해수문	A	“未發以扶正氣爲主 已發以散邪爲主”	COU-M40	발병 전후 분기
해수문	A	“冬風寒重 宜解表行痰…發散之後 以半夏等藥逐去其痰”	COU-M43	계절별 순차 개입
곽란문	B	“急用吐法” / “시간성은 주기가 아니라 속도”	GWR-SYS-01, 03	즉각 개입
곽란문	B	“霍亂吐瀉之時 切勿與穀食”	GWR-SYS-06	회복기 개입 유보
허로문	C	1 손~5 손 단계별 처치 분화 / “不受補者難治”	허-01~07, 허-49	단계적 분기
허로문	C	“峻補燥熱” 약 남용 금지	허-40	과잉 개입 경계
담음문	C	“先逐去敗痰 然後看虛實調理”	DM-08	순서적 개입
경계·정충문	D	“驚者平之” / “下視收神”	경계 NETWORK GB-009	행위를 통한 즉시 개입
경계·정충문	D	“若作痰治 往往殺人”	경계 NETWORK GB-010	오치 경계(개입 방향 오류 경고)
안문	D	“暴赤眼…不可點 亦不可洗”	ANM-SRC-21	급성기 개입 유보
소갈문	D	“消渴 病久 鍼灸를 금한다”	SGM-50~51	만성기 개입 유보

이 표에서 드러나는 타이밍 조율의 양상을 유형화하면 다섯 가지 패턴으로 정리된다.

첫째, 발병 전후 분기. 동일한 병증이라도 미발(未發)과 이발(已發)에서 치법이 정반대로 갈린다. 해수문의 “미발이면 부정, 이발이면 산사”가 대표적이다.

둘째, 단계적 분기. 병의 진행 단계에 따라 개입의 방법이 질적으로 달라진다. 허로문의 1손~5손 단계별 처치 분화, 소갈문의 상소→중소→하소 단계별 처방 분화가 이에 해당한다.

셋째, 즉각 개입. 급성 분기점에서 지체 없이 결정해야 하는 판단이다. 곽란문의 건곽란 응급 토법, 경계·정충문의 “驚者平之(놀란 자는 안정시킨다)”가 이에 해당한다.

넷째, 개입 유보. 적극적으로 ‘하지 않는 것’이 최선인 판단이다. 안문의 “급성 적안에 점안·세안 금지”, 광란문의 “토사 중 곡식 금지”, 소갈문의 “병이 오래되면 침뜸 금지”, 허로문의 “보약을 받아들이지 못하면 난치”가 이에 해당한다. 이 패턴은 특히 주목할 만 한데, 현대 의학에서는 ‘치료하지 않음’이 대개 포기의 의미를 갖지만, 조선 의학에서는 ‘지금은 손대지 않는 것’이 적극적이고 의식적인 판단으로 기록되어 있다.

다섯째, 오치 경계(개입 방향 오류 경고). 개입 자체는 필요하나 방향이 잘못되면 치명적이라는 경고다. 경계·정충문의 “若作痰治 往往殺人(만약 담으로 치료하면 종종 죽인다)”은 허손형 정충을 담병으로 오인하여 개입 방향을 잘못 잡는 것을 경고하는 서술이다. 허로문의 “峻補燥熱” 남용 금지도 같은 맥락으로, 성급하고 과격한 보법이 오히려 해가 됨을 명시한다.

이 특성을 본 연구는 ‘개입타이밍 조율의학(Intervention-Timing Medicine)’으로 명명한다. 개입타이밍 조율의학이란 치료적 개입의 시점·순서·강도를 병의 진행 단계와 전이 경로에 따라 신중하게 조율하며, 때로는 개입을 유보하거나 개입 방향의 오류를 경계하는 것 자체를 적극적 판단으로 포함하는 의학적 사유 방식을 뜻한다.

4.4.3 두 특성의 관계: 원인배열 위에서의 타이밍 조율

원인배열의학과 개입타이밍 조율의학은 별개의 특성이 아니라 상호 전제 관계에 있다. 복수의 원인이 배열되어 있기 때문에 어떤 원인에 먼저 개입할 것인지의 순서 문제가 발생하며, 이 순서의 결정이 곧 타이밍의 조율이다. 역으로, 개입의 타이밍을 조율하려면 현재 작동하고 있는 원인들의 배열 상태를 먼저 파악해야 한다.

이 관계를 데이터 안에서 확인할 수 있는 구체적 사례로 답음문을 들 수 있다. 답음문의 L2에서는 열→담과 수→음이라는 두 경로가 병렬적으로 배열되고, 경맥별로 열담·풍담·기담·한담이 분기된다(원인배열). 이 배열 상태를 파악한 위에서 L3의 “先逐去敗痰 然後看虛實調理”라는 순서적 개입이 결정된다(타이밍 조율). 만약 원인배열을 거치지 않고 일률적으로 담을 제거하면 허증 환자에게 해가 되며, 이것이 답음문이 “仲景三因의 표리·내외에 따른 구분에 비하면 곤담환 한 방으로 통치하는 것은 소략하다(DM-11)”고 자기 반성적으로 서술하는 이유다.

두 특성은 30병문 전체에 걸쳐 확인되나, 병문 유형에 따라 그 비중과 양상에 차이가 있다. 이를 유형별로 정리하면 다음과 같다.

표 15. 유형별 원인배열과 개입타이밍 조율의 양상 비교

유형	원인배열의 양상	타이밍 조율의 양상	대표 사례
A 생태순환형	몸-환경 상호작용 축 위에서 배열	계절·기후 변화와 연동된 조율	해수문: 가을 습→겨울 해수, 미발/이발 분기

B 판단노드형	배열 폭이 좁고 감별이 명확	‘즉각 개입 vs. 경과 관찰’ 이항 분기	곽란문: 전곽란=즉각 토법, 회복기=곡식 금지
C 네트워크허브형	다차원적 배열, 전이 경로 복수	가장 복잡한 조율, 순서·단계 모두 관여	허로문: 5 단계 분기+준보 금지+불수보 난치
D 경계판단형	내인/외인 감별이 일차 축	내부 기전 회복 경과에 따른 조율	안문: 급성 외치 금지, 경계문: 오치 경고

Type A에서는 원인배열이 몸-환경 상호작용의 축 위에서 이루어지며, 타이밍 조율이 계절·기후의 변화와 연동된다. 해수문의 “冬風寒重 宜解表行痰 發散之後 以半夏等藥逐去其痰(COU-M43)”은 겨울이라는 환경 조건 위에서 해표→행담→축담이라는 순차적 개입을 제시하는 서술로, 원인배열(계절-환경 축)과 타이밍 조율(순차 개입)이 하나의 판단 안에서 결합되어 있다.

Type B에서는 원인배열의 폭이 상대적으로 좁고 감별이 명확한 대신, 타이밍 조율은 ‘즉각 개입 vs. 경과 관찰’의 이항적 분기로 나타나는 경향이 있다. 곽란문에서 “시간성은 주기가 아니라 속도(GWR-SYS-01)”라는 메타 원리가 데이터에 명시적으로 기록되어 있는 것은, Type B 병문에서 타이밍이 판단의 핵심 변수임을 자기 서술적으로 보여준다.

Type C에서는 원인배열과 타이밍 조율 모두 가장 복잡한 양상을 보인다. 허로문의 경우 병원→오로→육극→칠상의 4단계 심화 구조(L2) 위에서 각 단계마다 다른 처치가 요구되고(L3), 동시에 “調和心腎 兼補脾胃(허-41)”라는 복수 장부 동시 개입의 원칙이 제시된다. 원인이 다차원적으로 배열되어 있으므로 타이밍 조율 역시 단일 축이 아니라 복수 축에서 동시에 이루어져야 하는 것이다.

Type D에서는 내인과 외인의 감별이 원인배열의 일차적 축을 이루며, 타이밍 조율은 내부 기전의 회복 경과에 따라 결정되는 경향이 있다. 안문의 “暴赤眼 不可點 亦不可洗(ANM-SRC-21)”는 급성 외감 단계에서 외치(外治)를 유보하고 내부 기전의 경과를 먼저 살피라는 판단이며, 경계·정충문의 “若作痰治 往往殺人”은 내인(심허)과 외인(담울)의 감별을 잘못했을 때의 치명적 결과를 경고하는 것으로, 원인배열의 정확성이 타이밍 조율의 전제 조건임을 극적으로 보여주는 사례라 하겠다.

5. 한계점 및 보완 방향: 관련 연구와의 비교

5.1 국내 한의학 디지털화 프로젝트와의 차이

한국한의학연구원의 〈한의학고전DB〉는 수백 종의 한의학 고전 원문과 번역을 데이터베이스화하여 웹으로 제공하는 국내 최대의 한의학 디지털 인프라다. 이 데이터베이스가 제공하는 핵심적 기능은 텍스트 접근성에 있다. 연구자는 〈한의학고전DB〉를 통해 『동의보감』, 『제중신편』 등의 원문과 번역문에 비교적 용이하게 접근할 수 있다.

한편으로 〈한의학고전DB〉의 데이터 스키마는 문헌 단위의 텍스트 제공에 특화되어 있으며, 병문 내부의 판단 구조를 층위별로 분절하거나, 병문 간의 관계를 정밀한 데이터 모델로 정의하는 것은 그 설계 목적에 포함되어 있지 않다. 본 연구에서 구축한 연구 데이터셋은 〈한의학고전DB〉가 제공하는 텍스트와 연계가 가능하도록 하되, 그 텍스트의 내부 구조를 판단 단위로 분절하고 L1·L2·L3 층위, 병문 유형, NETWORK 관계로 재조직한다는 점에서 기능적 층위가 다르다. 양자는 대체 관계가 아니라 보완 관계에 있다.

5.2 해외 전통의학 디지털화 프로젝트와의 비교

해외에서는 전통의학 지식의 구조화가 다양한 성격의 과업을 통해 진행되고 있다. 본 연구의 방법론적 위치를 명확히 하기 위해, 대표적 프로젝트들을 두 갈래의 축(구조화의 대상, 문헌 계보의 보존)에 근거하여 정리하였다.

5.2.1 TCM 지식그래프

중국의 TCM 지식그래프²⁰는 현재 세계에서 가장 방대한 전통의학 디지털 인프라다. OpenTCM은 48,000개 이상의 개체와 152,000개 이상의 관계로 구성된 다중 관계 지식 그래프를 구축하고 GraphRAG 기반 LLM과 통합하였다. 이 외에도 다수의 TCM 지식그래프 프로젝트가 빠르게 확장되고 있으며, LLM과 결합한 프로젝트도 보고되고 있다.

데이터 모델링 방법론의 관점에서 TCM 지식그래프와 본 연구의 차이는 다음과 같다. 첫째, 구조화 대상의 차이다. TCM 지식그래프는 약재·질병·증상·처방이라는 명사적 엔티티를 추출하고 그 사이의 관계를 매핑하는 것을 기본 방법론으로 삼는다. 구축의 핵심 작업은 개체명 인식(NER)과 관계 추출(RE)이다. 반면 본 연구는 엔티티의 추출이 아니라 판단 흐름의 분절을 기본 작업으로 삼는다. 예를 들어 “이 병증에 관련된 약제가 무엇인가”가 아니라 “이 병증을 어떤 순서로 인식하고, 어떤 기준으로 감별하며, 어떤 시점에 개입하는가”를 데이터화하는 것이다. 엔티티 중심의 구조화가 명사적(nominal) 접근

이라면, 본 연구는 절차적(procedural) 접근에 해당한다고 하겠다. 둘째, 문헌 계보 보존의 차이다. 대부분의 TCM 지식그래프는 현대에 표준화된 중의학 교과서나 임상 데이터를 지식의 원천으로 삼으며, 특정 역사적 문헌의 인용 계보를 데이터 구조 안에 보존하지 않는다. 본 연구의 Provenance_Citation 설계는 각 데이터 행이 어떤 문헌의 어떤 구절에 근거하는지를 기록하고, 조선 문헌이 인용한 중국 의서의 계보까지 추적함으로써 지식의 전송 경로를 데이터 안에 내장한다. 셋째, 지식 원천의 시간적 깊이 차이다. TCM 지식그래프의 지식 기반은 대부분 현대 표준화 시점에서 출발한다. 반면 본 연구의 데이터는 15세기(의방유취)에서 19세기(인제지)에 이르는 약 400년의 시간적 깊이를 가지며, 인용 계보를 통해 그 이전 시기의 중국 의서까지 소급된다. 이상의 비교를 정리하면 다음과 같다.

표 16. 데이터 모델링 방법론 비교: TCM 지식그래프 vs. 조선 의학 RDM(본 연구)

비교 항목	TCM 지식그래프	조선 의학 RDM(본 연구)
구조화의 대상	약재·질병·증상·처방 (명사적 엔티티)	판단 흐름 L1→L2→L3 (절차적 구조)
핵심 구축 방법	개체명 인식(NER) + 관계 추출(RE)	원문 분절 + 판단 층위 배정(수작업)
지식의 원천	현대 표준화 교과서·임상 데이터	역사적 문헌 3 축+ 1(15c-19c)
문헌 계보 보존	미보존 또는 단일 출처	Provenance_Citation 으로 계보 내장
오류 식별 기준	외부 검증 필요	텍스트 간 교차 대조로 내부 식별 가능
국가 플랫폼 등록	대부분 연구 프로토타입	DataON 공식 등록(DOI 부여)

다만 이 비교는 우열의 판단이 아니라 방법론적 지향의 차이를 보여주는 것임을 밝히는 바이다. TCM 지식그래프는 대규모 엔티티 네트워크를 통한 AI 응용(처방 추천, 진단 보조 등)에 강점이 있으며, 본 연구에서 구축한 데이터셋은 특정 시공간의 판단 구조를 정밀하게 재현하는 데 강점이 있다. 양자는 서로 다른 맥락의 질문에 대한 접근이라 할 수 있다.

5.2.2 『상한론(傷寒論)』 지식그래프 프로젝트

본 연구와 방법론적으로 가장 가까운 선행 연구는 중국에서 수행된 『상한론』 지식그래프 프로젝트들이다. 『상한론』을 대상으로 한 지식그래프 프로젝트로는, ‘병·맥·증·치’ 진단 사유 모델에 기반한 수작업 지식그래프 구축(Li), 지식 요소의 수작업 주석 체계 설계(Yang), 온톨로지 7단계 방법론을 적용한 상한론 온톨로지 구축(Zhang) 등이 보고되었다(Xiang et al. 2025에서 재인용²¹).

이 프로젝트들은 고전 텍스트 자체를 대상으로 삼고, 자동 추출이 아닌 수작업에 기초하며, 텍스트 내부의 진단 논리를 구조화하려 했다는 점에서 본 연구의 문제의식과 그 궤가 유사하다. 한편으로 차이도 있으니, 두 가지 갈래로 정리해볼 수 있다. 첫째, 『상한론』 프로젝트들은 단일 문헌을 대상으로 한다. 『상한론』이라는 하나의 텍스트 안에서 개체와 관계를 추출하는 방식이며, 복수의 문헌에 걸친 동일 병증의 판단 구조를 교차 분석하는 설계는 시도되지 않았다. 본 연구는 3축+1이라는 복수 문헌 체계를 설정하고, 동일 병문에 대해 『의방유취』, 『동의보감』, 『인제지』가 어떻게 다르게 서술되어 있는지를 Provenance_Citation을 통해 데이터로 비교할 수 있도록 설계하였다. 둘째, 『상한론』 프로젝트들의 구조화 단위는 여전히 엔티티(병·맥·증·치)에 기초한다. 엔티티 간의 관계가 진단 논리를 반영하고 있으나, 판단의 시간적 흐름(어떤 순서로 인식하고 감별하며 개입하는가)이 $L1 \rightarrow L2 \rightarrow L3$ 과 같은 층위 구조로 명시적으로 분절되어 있지는 않다. 이 점이 본 연구의 절차적 층위 모델과 구별되는 지점이라 하겠다.

5.2.3 인도의 TKDL 과 GRAYU

인도의 Traditional Knowledge Digital Library(TKDL)²²는 아유르베다, 유나니, 싯다, 요가의 전통 처방 45만 건 이상을 독자적 분류 체계(TKRC)로 구조화하여 국제특허분류(IPC)와 연동한 국가적 차원의 프로젝트다. TKDL의 학술적 기여는 전통지식을 국가 차원에서 체계적으로 분류·등록하여 생물해적행위(biopiracy)를 방지하는 법적·제도적 인프라를 구축한 데 있다. 본 연구의 DataON 등록이 연구데이터의 공공 자산화를 지향한다는 측면에서, 그러한 문제의식의 일부를 공유한다고 할 수 있겠다. 그러나 TKDL의 데이터 모델은 처방(formulation) 단위의 목록화에 특화되어 있다. TKDL이 “무엇이 처방되었는가?”라는 질문에 대한 구조화 작업이라면, 본 연구는 “왜, 어떤 판단을 거쳐 그 처방에 도달하는가?”라는 질문에 대한 구조화 작업이라 할 수 있다.

2025년 발표된 GRAYU²³는 아유르베다의 약용식물 12,000종 이상, 처방 1,000건 이상, 식물화학성분 약 130,000개, 질병 13,000개 이상을 그래프 데이터베이스로 통합한 프로젝트다. GRAYU의 방법론적 특징은 산스크리트 질병 용어를 MeSH·DOID 등 현대생의학 온톨로지와 매핑하여 전통의학과 현대의학 사이의 상호운용성(interoperability)을 확보한 점에 있다. 이는 본 연구가 아직 시도하지 않은 영역이며, 향후 조선 의학 연구 데이터셋이 지식그래프로 확장될 때 참고할 만한 연구 사례라 하겠다. 다만 GRAYU 역시 식물·화학성분·처방·질병이라는 명사적 엔티티를 구조화의 대상으로 삼고 있으며, 고전 텍스트 내부의 판단 흐름을 분절하는 접근은 취하고 있지 않다.

5.2.4 동남아시아 전통의학 디지털화

싱가포르의 Polyglot Asian Medicine Knowledge Graph²⁴는 동남아시아 전통의학의 다

언어 약재 지식을 Neo4j 그래프 데이터베이스로 구현한 프로젝트다. 약재명을 원어·영어·학명·문헌 출처·지리적 분포와 연결하는 구조로, 디지털 인문학 연구로서 전통 의학 지식을 그래프로 구현했다는 점에서 본 연구와 방법론적 친연성이 있다. 특히 다언어·다문화 맥락에서 지식의 교차와 번역을 추적한다는 문제의식은, 본 연구가 조선 문헌과 중국 문헌 사이의 인용 계보를 추적하는 것과 유사한 관심사에서 출발한다. 말레이 의학 필사본 디지털화 연구²⁵는 18~19세기 말레이 의학 필사본을 Protégé 기반 온톨로지로 구조화한 사례다. 비서구권 전통의학 필사본을 온톨로지 방법론으로 구조화한 점에서 본 연구와 비교할 만한 점이 있으나, 약재 분류와 치료 적용의 엔티티 구조화에 초점을 두었다는 측면에서 차이점 또한 분명하다.

5.2.5 종합 비교

이상의 프로젝트들을 두 축에 따라 정리하면 다음과 같다.

표 17. 해외 유관 프로젝트와의 종합 비교

프로젝트	구조화 대상	문헌 계보 보존	대상 문헌 수	판단 흐름 분절
TCM KG (OpenTCM 등)	명사적 엔티티	미보존	다수 (현대 표준)	없음
상한론 KG	엔티티(병·맥·증·치)	단일 문헌 내	1	부분적
TKDL	처방 목록	출처 기록 (계보 미추적)	다수	없음
GRAYU	명사적 엔티티	출처 기록 (계보 미추적)	다수	없음
Polyglot Asian Medicine	약재 엔티티 (다언어)	문헌 출처 연결	다수	없음
말레이 필사본 온톨로지	약재·치료 엔티티	필사본 단위	소수	없음
본 연구 (조선 의학 RDM)	절차적 판단 흐름	Provenance_Citation 계보 내장	3 축+1(4 종)	L1·L2·L3 층위 분절

이 표에서 드러나듯, 기존 프로젝트들은 “무엇이(what) 있는가?”를 구조화하는 데 초점이 있고, 본 연구는 “어떻게(how) 판단하는가?”를 구조화하는 데 초점이 있다. 또한 문헌 계보의 보존에 있어서, 기존 프로젝트들이 출처를 기록하는 수준에 머무는 반면 본 연구는 문헌 간 인용의 전승 경로를 데이터 구조 안에 내장하여 교차 대조와 오류 식별이 가능하도록 설계하였다. 이 두 가지가 여타 선행 연구들과 비교할 때 본 연구의 데이터 모델링 방법론이 갖는 뚜렷한 차별점이라 하겠다.

5.3 본 연구의 한계와 후속 연구의 방향

본 연구의 방법론에는 몇 가지 뚜렷한 한계가 있다. 후속 연구를 위해 그에 관한 내용을 정리할 필요가 있다.

첫째, 규모의 한계다. 『의방유취』의 91개 병문 가운데 30개만이 1차 대상으로 구축되었다. 30병문은 조선 의학 판단 체계의 전모를 대표하기에 충분한 규모라고 단정하기 어렵다. 4장에서 논의한 유형별 분포와 구조적 특성이 나머지 61개 병문에서도 동일하게 확인되는지는 향후 확장 작업을 통해 검증되어야 한다. 특히 현재 포함되지 않은 병문 가운데 Type A-D의 어느 유형에도 깔끔하게 분류되지 않는 사례가 출현할 가능성이 있으며, 이 경우 유형 분류 체계 자체의 수정이 필요할 수 있다. 본 30병문은 본 연구의 종착점이 아니라 전체 병문(약 50병문 예상)으로 나아가는 과정의 단계적 한 단면이다.

둘째, 해석적 재현성과 다해석 병렬 수용 설계의 과제다. 본 프로젝트의 데이터 구축은 1인 연구자에 의해 수행되었으며, 원문을 판단 단위로 분절하고 L1·L2·L3 층위를 배정하는 과정에는 불가피하게 연구자의 해석적 판단이 개입한다. 또한 한문 텍스트는 그 본래적 의미론적 다층성으로 인하여, 동일한 구절에 대해 연구자마다 다소 다른 해석이 가능하다. 본 연구는 ‘문헌 우선 원칙’, ‘분리 원칙’, ‘문장이 스스로 말하게 하라’는 세 가지 운용 원칙을 통해 연구자의 자의적 개입을 최소화하고, 또한 3.3절의 L1·L2·L3 판별 기준표를 통해 층위 배정의 운용 기준을 명시함으로써 후속 연구자의 재현 가능성을 확보하고자 하였으나, 이러한 장치들이 한문 해석의 주관성 자체를 해소한다고 보기는 어렵다.

이 한계의 본질은 단순히 ‘코더 간 신뢰도 검증의 부재’로 환원되지 않는다. 한문 텍스트의 해석은 다수의 연구자가 같은 답에 수렴해야 하는 객관적 작업이 아니라, 텍스트의 의미론적 다층성을 학술 공동체가 누적적으로 협상하고 재해석하는 과정이다. 따라서 본 연구가 직면하는 과제는 ‘올바른 단일 해석의 발견’이 아니라 ‘복수의 해석을 데이터 구조 안에 어떻게 수용할 것인가’이다. 본 연구는 이 인식 위에서, 본 RDM의 후속 단계 확장 설계로서 다음과 같은 다해석 병렬 수용 스키마(Pluralistic Interpretation Schema)를 제시한다.

핵심 설계는 RDM_Core 스키마에 두 개의 필드를 도입하는 것이다. 첫째, Interpreter_ID 필드는 각 데이터 행의 해석 주체를 식별한다. 둘째, Interpretation_Variant 필드는 동일 원문 행에 대한 대안 해석을 동일 데이터셋 안에 병렬로 수록한다. 이 두 필드의 도입으로 본 RDM은 ‘이 원문 행은 연구자 A에 의해 L2(판단 구조)로 배정되었으나, 연구자 B는 이를 L3(치료 실행)으로 본다’는 식의 해석적 다양성을 데이터 구조 자체 안에서 보존·비교할 수 있게 된다. 곧 본 RDM은 단일 정답의 데이터셋이 아니라, 학술 공동체의 누적 검증과 재해석을 통해 진화하는 열린 데이터 자산으로 자기 정체성을 재정립한다.

이러한 설계 방향은 디지털 인문학이 직면해 온 ‘주관성 대 객관성’의 이분법을 우회하

여, 해석의 다층성 자체를 데이터의 구조적 자산으로 전환한다. 이는 본 논문 3.3절의 판별 기준표(해석의 운용적 정밀화)와 짝을 이루어, 본 RDM의 재현 가능성을 두 축에서 보장한다. 한 축은 운용 기준의 명시화를 통한 해석의 일관성 제고이고, 다른 한 축은 다해석 병렬 수용을 통한 해석의 다양성 보존이다. 두 축은 상호 대립하지 않고, 한문 텍스트의 본래적 다층성을 데이터 모델 안에 정직하게 반영하는 두 가지 보완적 장치라 하겠다. 다해석 병렬 수용 스키마의 구체적 구현, 곧 본 RDM의 후속 버전으로의 적용은 본 연구의 후속 과제이며, 본 연구의 학술적 기여 중 하나인 ‘다해석 병렬 수용 스키마 설계의 단초 제공’으로서 7장 결론에서 명시한다.²⁶

셋째, 『역시만필』의 ‘+1’ 활용 범위의 한계다. 『역시만필』의 130개 의안이 30병문 전체를 균등하게 포괄하지는 않는다. 실제 데이터를 확인한 결과, 30병문 중 11개 병문(소갈문, 황달문, 객란문, 대변문, 적취문, 해수문, 창만문, 구토문, 제기문, 혈병문, 허로문)에 대해 『역시만필』의 관련 의안 또는 임상 사례가 데이터셋 안에 반영되어 있으며, 나머지 19개 병문에 대해서는 직접 대응하는 데이터가 현재 포함되어 있지 않다. 이 중 소갈문과 황달문은 『역시만필』 의안이 STAGE3 시트로 별도 구축되어 임상 교차 검증이 가장 두텁게 이루어진 사례이고, 객란문·대변문·적취문도 다수의 의안이 반영되어 있다. 반면 부종문, 해학문, 담음문, 내상문, 온역문, 안문, 비문, 이문, 치문, 소변문, 전간광문, 경계·정충·건망문, 조열문, 중한·고랭문, 습병문, 해독문, 금창문, 중서문, 제상문 등 19개 병문에서는 『역시만필』과의 교차 검증이 이루어지지 않았다. ‘+1’ 체계의 실제적 유효 범위는 현재 약 37%(11/30)에 해당하며, 이 범위의 확장이 후속 과제라 하겠다.

넷째, NETWORK 관계 어휘의 비표준화 문제다. 4.3절에서 확인된 바와 같이, 30병문의 NETWORK 시트에서 사용된 고유 관계 유형은 37종에 달한다. 이 중 SPLITS_TO, BRIDGES_TO, PART_OF 등 15개 이상의 병문에서 공통적으로 사용되는 유형이 있는 반면, IS_YANG_FORM, PAIRS_WITH, ASSESSED_BY 등 1개 병문에서만 사용되는 유형도 다수 존재한다. 이는 데이터 구축이 30병문에 걸쳐 장기간 진행되면서 관계 어휘가 점진적으로 확장된 결과다. 병문별 텍스트의 서술 특성에 따라 새로운 관계 유형이 귀납적으로 추가된 것은 ‘문장이 스스로 말하게 하라’는 원칙의 자연스러운 귀결이나, 동시에 병문 간 관계 데이터의 통합적 분석을 어렵게 만드는 요인이기도 하다. 의미적으로 유사한 관계 유형(예: LEADS_TO와 TRANSITIONS_TO, BRIDGES_TO와 LINKS_TO)의 정리·표준화는 향후 지식그래프 확장 단계에서 반드시 선행되어야 할 작업이라 하겠다.

상기한 바 여러 한계에도 불구하고, 본 연구는 네 가지 측면에서 뚜렷한 학술적 기여가 있다고 판단된다. 첫째, 전통의학 고전 텍스트의 판단 흐름을 절차적으로 분절하여 데이터셋으로 구축한 최초의 체계적 시도라는 점이다. 둘째, 복수 문헌의 인용 계보를 데이터 구조 안에 내장하는 Provenance_Citation 설계를 제시한 점이다. 『상한론』 지식그래프를 포함한 기존 프로젝트 중 어느 것도 문헌 간 인용의 전송 경로를 데이터 구조 안에서 추적 가능하게 설계하지는 않았다. 셋째, 구축된 데이터셋을 국가 연구데이터 플랫폼에 공식 등록하여 재현 가능하고 검증 가능한 공공 자산으로 전환한 점이다. 넷

재, 한문 텍스트의 본래적 다해석성을 데이터 구조의 결함이 아니라 구조적 자산으로 전환하는, 다해석 병렬 수용 스키마(Pluralistic Interpretation Schema) 설계의 단초를 제시한 점이다. 앞서 밝힌 4가지 한계점은 이러한 학술적 기여를 부정하는 것이 아니라, 향후 후속 연구를 진행할 시 반드시 보강해야 할 구체적 방향을 제시한 것이라 하겠다.

6. 지식그래프 확장과 AI 연계의 전망

이 장에서는 30병문 데이터셋을 기반으로 한 지식그래프 확장의 설계 방향, AI 학습 및 참조 데이터로서 활용 가능성과 그 조건, 그리고 전통 지식 데이터의 공공 자산화 원칙에 대해 논의한다.

6.1 3층 지식그래프로의 확장 설계

현재 구축된 30병문 데이터셋은 각 병문 내부의 판단 구조(RDM_Core 약 750행)와 병문 간 관계(NETWORK 약 485건)를 담고 있다. 이 데이터셋은 그 자체로 이미 완결된 연구데이터라고 할 수 있겠으나, 추후 보완이 이루어질 경우 지식그래프(Knowledge Graph)로 확장될 가능성을 지니고 있다. 확장의 설계는 3층 구조를 따른다.

표 18. 지식그래프 확장의 3층 설계

층위	그래프 명칭	노드 유형	데이터 원천
1층	Disease Graph	병문, 증상, 감별 기준	RDM_Core L1·L2 데이터
2층	Formula Graph	처방, 약재 조합, 적응증	RDM_Core L3 + FORMULA_LAYER
3층	Citation Lineage Network	문헌, 인용 관계, 전승 경로	Provenance_Citation 데이터

1층 Disease Graph는 병문과 증상, 병문 간 관계를 노드와 엣지로 구현한다. 4장에서 분석한 NETWORK 관계 데이터 즉, 전이·분기(약 19%), 연결·교량(약 19%), 인과·생성(약 14%) 등이 이 그래프의 직접적 기반이 된다. 4.3절에서 확인된 허로문(46건), 제기문(30건), 혈병문(30건) 등 허브 병문의 위상이 그래프 구조 안에서 시각적으로 드러나게 된다.

2층 Formula Graph는 각 병문의 L3 데이터에서 처방과 약재 조합을 추출하여, 병문에서 처방을 호출하고 처방이 약재 조합으로 구현되는 치료 알고리즘을 그래프로 표현한다. 현재 일부 병문(허로문, 소갈문, 경계·정충문 등)에는 이미 FORMULA_LAYER 시트가 별도로 구축되어 있어, HAS_FORMULA 관계(陰虛→大補陰丸, 陽虛→桂附湯, 上消→白虎加人蔘湯 등)가 2층 그래프의 초기 데이터로 직접 활용될 수 있다.

3층 Citation Lineage Network는 Provenance_Citation 데이터를 기반으로 문헌 간 인용의 전송 경로를 시각화한다. 해수문의 경우 『河間方』, 『內經』, 『丹溪心法』, 『明理論』, 『金匱方』 등 다수의 중국 의서가 Provenance_Citation으로 기록되어 있어, ‘조선 편찬서→인용된 중국 원전’의 계보가 데이터를 통해 추적 가능하다. 이 3층이 통합되면 병문(Disease)-처방(Formula)-약재(Herb)-문헌(Provenance)을 별도 노드로 분리하고, 노드 유형(Type A-D)과 엣지 온톨로지를 속성으로 표준화하는 하나의 탐색 가능한 지식 네트워크가 구현된다.

이 3층 구조 위에 후속 단계로 ‘해석자 층(Interpreter Layer)’의 도입이 가능하다. 해석자 층은 그래프의 별도 노드·엣지 차원이 아니라, 1층 Disease Graph의 각 노드·엣지에 부착되는 메타 데이터의 층위로 설계된다. 곧 동일한 노드(예: 해수문)나 엣지(예: 담음→해수)에 대한 복수 연구자의 해석적 판단이 동일 그래프 안에 병렬로 보존되며, 그 그래프 탐색 시 ‘누구의 해석 위에 본 결과가 서 있는가’를 명시적으로 추적할 수 있다. 이는 5.3절 둘째 항목에서 제시한 RDM_Core 스키마의 Interpreter_ID·Interpretation_Variant 필드 설계를 그래프 차원으로 확장한 결과이며, 본 지식그래프가 단일 해석의 권위적 표상이 아니라 학술 공동체의 협업적 해석 누적 가능한 열린 구조임을 의미한다. 해석자 층의 구체적 구현은 본 연구의 후속 과제이나, 그 설계의 핵심 원칙은 미리 명시할 수 있다. (가) 동일 노드·엣지에 대한 복수 해석은 충돌이 아니라 자산으로 보존된다. (나) 각 해석에는 해석 주체(Interpreter_ID), 해석 근거(원문 인용), 해석 시점(timestamp)이 메타 데이터로 부착된다. (다) 그래프 탐색 시 사용자는 특정 해석 주체의 시점에서 본 그래프 형태를 필터링하여 조회할 수 있으며, 동시에 복수 해석의 중첩 형태로도 조회할 수 있다.

다만 5.3절에서 밝힌 바와 같이, 현재 37종에 달하는 NETWORK 관계 어휘의 표준화가 지식그래프 확장의 선행 과제다. SPLITS_TO와 TRANSITIONS_TO, BRIDGES_TO와 LINKS_TO처럼 의미적으로 중첩되는 관계 유형을 정리하고, 병문 전체에 걸쳐 일관된 엣지 온톨로지를 확립해야 한다. 현재 30병문 규모에서 향후 『의방유취』의 91개 병문 전체로 확장되면 노드와 엣지의 수는 선형이 아닌 지수적으로 증가하며, 지식그래프의 분석적 유용성도 그에 비례하여 높아질 것으로 예상된다.

6.2 AI 활용 데이터로서 가능성과 조건

30병문 데이터셋이 대규모 언어모델(LLM)이나 그래프 기반 추론 시스템의 학습 또는 참조 데이터로 활용되려면, 몇 가지 구체적 조건이 충족되어야 한다.

첫째, 데이터 규모의 문제다. 현재 30병문 데이터셋 규모는 LLM의 사전 학습(pre-training)에 직접 활용하기에는 작다. 그러나 특정 도메인의 구조화된 지식을 LLM에 주입하는 방식으로는 사전 학습보다 검색증강생성(RAG)이나 파인튜닝(fine-tuning)이 더 적합하며, 이 방식에서는 데이터의 양보다 구조의 정밀성과 출처의 추적 가능성이 더 중

요하다. 본 연구의 RDM은 L1·L2·L3 층위와 Provenance_Citation이 내장되어 있어, RAG 시스템에서 “어떤 문헌의 어떤 판단에 근거하여 이 답변을 생성했는가?”를 추적할 수 있는 구조를 갖추고 있다. 예컨대 해수문의 COU-M40에 대해 “『동의보감』 喘證 총론, Provenance: 『丹溪心法』”이라는 출처 정보가 내장되어 있으므로, AI 시스템이 “미발 시 부정(扶正), 이발 시 산사(散邪)”라는 판단을 생성할 때 그 근거를 명시적으로 제시할 수 있다. 이는 TCM 지식그래프 기반 LLM 시스템에서 지적되는 블랙박스 문제(판단의 근거를 추적할 수 없는 문제)에 대한 하나의 방법론적 대안이 될 수 있다.

둘째, 표준화 수준의 문제다. 현재 데이터셋의 용어 체계는 조선 의학 원전의 용어를 그대로 보존하는 방식으로 설계되어 있으며, 현대 의학 용어나 국제 표준 온톨로지(ICD, MeSH, SNOMED CT 등)와의 매핑은 이루어지지 않았다. 이 매핑이 없으면 AI 시스템이 조선 의학 데이터를 현대 의학 데이터와 연계하여 활용하기 어렵다. 5.2.3항에서 언급한 GRAYU의 산스크리트-MeSH/DOID 매핑 사례²⁷는 이 과제의 구체적 참조점이 된다. 다만 이 매핑은 단순한 용어 대응이 아니라 서로 다른 의학 체계의 개념적 차이를 고려한 신중한 작업이 필요하며, 본 연구의 범위를 넘어서는 후속 과제다. 특히 4.4절에서 확인된 원인배열의학의 특성(단일 원인 환원을 거부하고 복수 원인을 배열하는 사유 방식)은 현대 의학의 단일 진단 코드 체계와 근본적으로 다른 구조를 가지므로, 매핑 작업에는 이 차이에 대한 방법론적 고려가 전제되어야 할 것이다.

셋째, 라이선스와 접근성의 문제다. 현재 DataON에 등록된 데이터셋은 공공 연구데이터로서 접근이 가능하나, AI 학습 및 참조에 활용하기 위해서는 라이선스 조건을 명확히 할 필요가 있다. 현재 DataON 등록 시점에서 라이선스 조건이 명시적으로 부여되지 않았으며, 이는 AI 활용을 위해 향후 정비가 필요한 사항이다.

이상의 조건을 종합하면, 30병문 데이터셋이 AI에 제공할 수 있는 고유한 가치는 대규모 데이터로서의 양적 기여가 아니라, 판단의 구조와 근거가 추적 가능한 형태로 정리된 질적 기여에 있다. 4장의 분석에서 L2(판단 구조)가 전체 데이터의 약 45%를 차지한다는 사실은, 이 데이터셋이 단순한 증상-처방 매핑이 아니라 판단의 논리 자체를 기록하고 있음을 보여준다. AI 시대에 전통의학 데이터가 갖는 의미는 방대한 양의 텍스트를 학습시키는 데 있는 것이 아니라, 인간의 판단 구조를 명시적으로 분절하고 그 근거를 투명하게 제시하는 데이터를 제공하는 데 있을 수 있다. 조선 의학이 보여주는 판단의 특성(복합 원인을 배열하고, 개입의 타이밍을 조율하며, 때로는 개입을 유보하는 판단)은 AI의 자동 추론만으로 온전히 수용하여 장착하기는 어려운 영역이기도 하다.

6.3 전통 지식 데이터의 FAIR·CARE 원칙 적용

본 연구의 데이터셋이 DataON에 등록되었다는 사실은 곧 FAIR 원칙의 구체적 실천을 의미한다. FAIR 원칙²⁸ 즉, Findable(검색 가능), Accessible(접근 가능), Interoperable(상호운용 가능), Reusable(재사용 가능)은 과학 데이터의 관리와 공유를 위한 국제적 가이

드라인으로, 최근 인문학 분야에서도 그 적용이 논의되고 있다. 30병문 데이터셋의 FAIR 충족 현황을 점검하면 다음과 같다.

- Findable: DataON 플랫폼에 등록되어 검색이 가능하며, DOI가 부여되었다.
- Accessible: 플랫폼을 통해 접근이 가능하다.
- Interoperable: 현재는 조선 의학 내부의 용어 체계를 사용하고 있어, 타 분야·타 체계와의 상호운용성은 제한적이다. 5.3절에서 지적한 37종의 NETWORK 관계 어휘 비표준화 문제와 6.2절에서 논의한 국제 표준 온톨로지 매핑의 부재가 이 제한의 구체적 내용이다.
- Reusable: 데이터셋 스키마와 설계 원칙이 「조선의학 사유구조 RDM 프로젝트 헌법」으로 공개되어 있어, 다른 연구자가 이를 기반으로 추가 병문을 구축하거나 검증 작업을 수행할 수 있다. 특히 ‘분리 원칙’에 따라 한문 원문과 한글 번역문이 별도 필드로 보존되어 있어, 해석의 타당성을 독립적으로 검토하는 것이 가능하다.

한편 전통 지식 데이터에는 FAIR 원칙만으로는 충분하지 않은 측면이 있다. Carroll 등이 제안한 CARE 원칙²⁹ 즉, Collective Benefit(집합적 이익), Authority to Control(통제 권한), Responsibility(책임), Ethics(윤리)는 토착 지식과 전통 지식의 데이터 거버넌스를 위한 보완적 프레임워크다. 조선 의학 데이터셋은 특정 토착 공동체의 지식이 아니라 역사적 문헌에 기초한 학술 데이터이므로 CARE 원칙의 직접적 적용 대상은 아니다. 그러나 전통 지식이 AI 학습 및 참조 데이터로 활용될 때 그 지식의 맥락과 출처가 보존되어야 한다는 CARE의 핵심 문제의식은 본 연구의 Provenance_Citation 설계와 맥을 같이한다. 해수문의 56행 데이터 각각에 『河間方』, 『內經』, 『丹溪心法』, 『金匱方』 등의 출처가 기록되어 있는 것은, 이 데이터가 맥락 없이 유통되는 것을 구조적으로 방지하는 장치이기도 하다.

최근 국가의 자체 AI 생산 역량과 데이터 주권과 관련해서 ‘소버린 AI(Sovereign AI)’에 관한 국제적 논의가 확대되고 있다. 이 논의의 핵심 질문 중 하나는 ‘누구의 지식 체계가 AI에 내장되는가’이다. 5.2.1항에서 비교한 바와 같이, 중국의 TCM 지식그래프가 자국의 전통의학 개념을 글로벌 AI 인프라에 탑재하는 작업이라면, 조선 의학 데이터셋은 동일한 동아시아 의학 지식의 자장을 공유하면서도 그것을 어떻게 다르게 판단하고 적용했는가를 데이터화한 것이다. 4.4절에서 귀납된 원인배열의학과 개입타이밍 조율의학은 TCM 지식그래프의 명사적 엔티티 구조로는 포착되지 않는, 조선 의학 고유의 절차적 판단 구조다. FAIR와 CARE를 통합하는 프레임워크 위에서, 각국의 고유한 판단 구조가 데이터로 구축되고 국제적으로 연결될 때, 다양한 인식론이 AI 안에 공존할 수 있는 토대가 마련될 수 있음을 주지할 필요가 있다.

7. 결론

본 연구는 조선 의학 고전문헌에 내재된 질병 판단 알고리즘을 데이터로 설계·구축하는 방법론을 제시하였다. 『의방유취』, 『동의보감』, 『인제지』를 축적·표준화·응용의 3축으로, 『역시만필』을 임상 검증 텍스트로 설정하는 ‘3축+1’ 문헌 체계 위에서, 30개 병문의 판단 흐름을 L1(현상 인식)·L2(판단 구조)·L3(치료 실행)의 3층위로 분절하고, 병문의 판단 기능에 따라 생태순환형(A)·판단노드형(B)·네트워크허브형(C)·경계판단형(D)의 4유형으로 분류하였다. 각 데이터 행에 Provenance_Citation을 내장하여 문헌 간 인용 계보를 데이터 구조 안에서 추적할 수 있도록 설계하였으며, 구축된 데이터셋(RDM_Core 약 750행, NETWORK 약 485건)은 KISTI DataON 국가연구데이터플랫폼에 공식 등록되었다. 편찬 데이터를 분석한 결과, L2(판단 구조)가 전체의 약 45%로 가장 높은 비율을 차지하여 조선 의학 문헌이 판단의 논리를 서술하는 데 가장 많은 지면을 할애하고 있음이 확인되었고, Type C(네트워크허브형)의 평균 NETWORK 엣지 수(22.1건)가 다른 유형(11.2~16.2건)에 비해 뚜렷하게 높아 유형 분류와 관계 데이터 사이의 정합성이 사후적으로 검증되었다.

이 구축 과정을 통해 조선 의학의 두 가지 구조적 특성이 귀납적으로 확인되었다. 병의 원인을 단일 요인으로 환원하지 않고 복수의 원인을 체계적으로 배열하여 그 관계 속에서 판단을 구성하는 ‘원인배열의학’과, 개입의 시점·순서·강도를 병의 진행 단계와 전이 경로에 따라 신중하게 조율하며 때로는 개입 자체를 유보하는 ‘개입타이밍 조율의학’이 그것이다. 30병문의 대표 사례 분석에서 원인배열의 축 수가 모든 병문에서 3 이상으로 확인되었고, 타이밍 조율의 양상은 발병 전후 분기, 단계적 분기, 즉각 개입, 개입 유보, 오치 경계의 다섯 가지 패턴으로 유형화되었다. 이 두 특성은 5장에서 비교한 TCM 지식그래프를 비롯한 해외 전통의학 디지털화 프로젝트들이 주로 명사적 엔티티의 추출에 초점을 두는 것과 방법론적으로 구별되는 지점이며, 본 연구가 판단의 절차적 흐름을 구조화 대상으로 삼았기에 가시화할 수 있었던 결과이기도 하다.

이상의 작업을 통해 본 연구는 다음 네 가지 측면에서 학술적 기여를 정리해 볼 수 있다. 첫째, 전통의학 고전 텍스트의 판단 흐름을 절차적으로 분절하여 공공 연구데이터로 구축한 최초의 체계적 시도라는 점이다. 둘째, 복수 문헌의 인용 계보를 데이터 구조 안에 내장하는 Provenance_Citation 설계를 통해, ‘이 판단이 어떤 문헌 전통에서 왔는가’가 데이터 차원에서 추적 가능한 구조를 제시한 점이다. 셋째, 구축된 데이터셋을 KISTI DataON 국가연구데이터플랫폼에 공식 등록하여 재현 가능하고 검증 가능한 공공 자산으로 전환한 점이다. 넷째, 한문 텍스트의 본래적 다해석성을 데이터 구조의 결함이나 구조적 자산으로 전환하는 다해석 병렬 수용 스키마(Pluralistic Interpretation Schema) 설계의 단초를 5.3절·6.1절에서 제시한 점이다.³⁰⁾ 이 네 번째 기여는 본 RDM이 단일 권위의 종결적 데이터셋이 아니라, 학술 공동체의 누적 검증과 재해석을 통해 진화하는 열린 데이터 자산임을 자기 정체성의 일부로 명문화한 것이며, 디지털 인문학

이 직면해 온 ‘주관성 대 객관성’의 이분법을 우회하는 하나의 방법론적 단서를 제공한다.

본 연구의 데이터셋은 3층 지식그래프(Disease Graph-Formula Graph-Citation Lineage Network)와 그 위에 부착되는 해석자 층(Interpreter Layer)으로의 확장과 RAG 기반 AI 시스템의 참조 데이터로서의 활용 가능성을 열어 준다. 그 고유한 가치는 대규모 데이터의 양적 기여가 아니라, 판단의 구조와 근거가 추적 가능한 형태로 정리된 질적 기여에 있다. 30병문이라는 규모, 1인 연구자 중심 구축의 재현성, 한문 해석의 주관성, 『역시만필』 교차 검증의 제한적 유효 범위(11/30, 약 37%), 37종에 달하는 NETWORK 관계 어휘의 비표준화 등 본 연구의 한계는 5장에서 명시한 바와 같으며, 이는 후속 연구를 통해 보강해야 할 구체적 과제라 하겠다. 전통의학 고전 텍스트의 판단 흐름을 절차적으로 분절하여 공공 연구데이터로 구축한 본 연구의 시도가, 디지털 인문학 방법론을 적용한 하나의 사례로서 후속 연구 및 학술적 논의의 구체적 단서가 되기를 바란다.

¹ 이 흐름의 주요 성과로 다음을 들 수 있다. 신동원 (2015). 『동의보감과 동아시아 의학사』. 한국의 과학과 문명 1. 들녘. 김호 (2000), 『허준의 동의보감 연구』, 일지사. 이경록 (2017), 「『의방유취(醫方類聚)』의 편찬과 그 함의」, 『국역 의방유취』 1, 세종대왕기념사업회. 박훈평·안상우 (2014), 「혜민서 관청지 『惠局志』 편제와 내용 연구」, 『한국의사학회지』 27(2), 한국의사학회. 등 관련 논저 다수.

² 『대한한의학원전학회지』(대한한의학원전학회 간), 『한국의사학회지』(한국의사학회 간). 두 학회는 1987 년 창간된 『원전의사학회지』에서 1999 년 분화된 것으로, 전자는 원전·이론 연구를, 후자는 의학사 연구를 중심으로 학술 활동을 이어 오고 있다.

³ 한국한의학연구원 한의학교전 DB. URL: <https://mediclassics.kr>. 한국전통지식포털(KoreanTK). URL: <https://www.koreantk.com>

⁴ 한의약융합연구정보센터(KMCRIC, URL: <https://www.kmcric.com>.)는 한약·침구·경혈·생약 성분·약물 상호작용 등에 관한 임상·실험 연구 정보를 근거중심의학(EBM) 방법론으로 재가공하여 제공해 왔고, OASIS 전통의학정보포털(URL: <https://oasis.kiom.re.kr>.) 역시 전통의학 학술 정보를 임상·연구 현장과 매개하는 역할을 수행하고 있다.

⁵ Wang, S. et al. (2024). "A Review of Knowledge Graph in Traditional Chinese Medicine." Computers, Materials & Continua, 81(3). ; Liu, C. et al. (2024). "Research on Traditional Chinese Medicine: Domain Knowledge Graph Completion and Quality Evaluation." JMIR Medical Informatics, 12, e55090. ; He, J., Guo, Y., Lam, L.K., Leung, W., He, L., Jiang, Y., Wang, C.C., Xing, G., & Chen, H. (2025). "OpenTCM: A GraphRAG-Empowered LLM-based System for Traditional Chinese Medicine Knowledge Retrieval and Diagnosis." arXiv:2504.20118.

⁶ Xiang, S. et al. (2025). "Integrating Knowledge Graphs with Ancient Chinese Medicine Classics: Challenges and Future Prospects of Multi-Agent System Convergence." Chinese Medicine 20, 168. <https://doi.org/10.1186/s13020-025-01226-7> -본 논문의 참고문헌 [21], [22], [23] 참조.

⁷ CSIR-TKDL. Traditional Knowledge Digital Library (URL: <https://www.csir.res.in/en/documents/tkdl>) ; Fredriksson, M. (2022). "India's Traditional Knowledge Digital Library and the Politics of Patent Classifications." Third World Quarterly 44(4), 813-832. <https://doi.org/10.1080/01436597.2022.2146529>

⁸ GRAYU (URL: <https://caps.ncbs.res.in/GRAYU/>) ; Joshi, S. et al. (2025). "GRAYU: Graph-based Database Integrating Ayurvedic Formulations, Medicinal Plants, Phytochemicals and

Diseases." *Frontiers in Pharmacology* 16, 1727224. <https://doi.org/10.3389/fphar.2025.1727224>

⁹ Polyglot Asian Medicine Knowledge Graph (URL: <https://kgraph.sg/polyglot/>) ; Khoo, C.S.G., Stanley-Baker, M., Zakaria, F.B., Chen, J., Ang, S.Q.R., & Huang, B. (2023). "Development of the Polyglot Asian Medicine Knowledge Graph System." ICADL 2023, LNCS 14458, 3–11. https://doi.org/10.1007/978-981-99-8088-8_1

¹⁰ Fauzi, N., Mohamed Najid, M.H., Ibrahim, N., & Sahbudin, M.A. (2025). "Digitalizing Malay Medical Manuscripts: A Database-Driven Approach to Preservation, Knowledge Management and Intellectual Property." *Journal of Al-Tamaddun* 20(2), 199–210. <https://doi.org/10.22452/JAT.vol20no2.14>

¹¹ 전종욱 (2025a). 「조선의학 사유구조 RDM 프로젝트 헌법」. 전북대학교 한국과학문명학연구소. 전종욱 (2025b–2026). 「조선의학 질병 인식구조 RDM — 30 병문」. KISTI DataON 국가연구데이터플랫폼 URL: <https://dataon.kisti.re.kr>

¹² 안상우 (2000). 『의방유취에 대한 의사학적 연구』. 경희대학교 박사학위논문.

¹³ 신동원 (2015). 『동의보감과 동아시아 의학사』. 한국의 과학과 문명 1. 들녘.

¹⁴ 전종욱·조창록, 「『임원경제지』·『인제지』의 편집 체제와 조선후기 의학 지식의 수용 양상: 『동의보감』과의 비교를 중심으로」, 『의사학』 21(3), 대한의사학회, 2012, 403–448 쪽.

¹⁵ 이수귀 저, 신동원·오재근·이기복·전종욱 공역(2015), 『역시만필: 조선어의 이수귀의 동의보감 실천기』. 들녘. 오재근 (2022), 「17–8 세기 조선의 임상 의학은 어떻게 이뤄졌을까 - 『역시만필』을 통해 살펴본 의관 이수귀의 임상 의학-」, 『의사학』 31(1) 70호, 대한의사학회, 1–34 쪽.

¹⁶ 전종욱 (2025). 「조선의학 사유구조 RDM 프로젝트 헌법」. 전북대학교 한국과학문명학연구소.

¹⁷ 차후 연구는 전체 병문을 모두 RDM 데이터셋으로 만드는 것을 목표로 하며, 완성된 병문의 개수는 50 개 정도로 예상된다. 지금 남아있는 병문들은 상한문(傷寒門) · 제풍문(諸風門) · 웅저문(癰疽門) · 부인문(婦人門) · 소아문(小兒門) 등 분량이 방대하여 하나의 병문만으로도 별도의 분석이 필요할 정도이므로 방안을 고려해야 한다. 예를 들어 상한문(傷寒門)은 병문 하나가 『의방유취』 전체 분량의 20%를 상회하는데, 최소한 6 경 병증을 나누어 별도 병문으로 처리해야 안정적 분석이 가능할 것이다.

¹⁸ 『의방유취』는 세종대왕기념사업회본, 「인제지」는 임원경제연구소본을 저본으로 삼았고, 『동의보감』은 <한의학고전 DB>의 번역을 원용했다. <https://dataon.kisti.re.kr>

¹⁹ 전종욱 (2025b–2026). 「조선의학 질병 인식구조 RDM — 30 병문」. KISTI DataON 국가연구데이터플랫폼. <https://dataon.kisti.re.kr>

²⁰ Wang, S. et al. (2024). "A Review of Knowledge Graph in Traditional Chinese Medicine." *Computers, Materials & Continua*, 81(3). ; Liu, C. et al. (2024). "Research on Traditional Chinese Medicine: Domain Knowledge Graph Completion and Quality Evaluation." *JMIR Medical Informatics*, 12, e55090. ; He, J., Guo, Y., Lam, L.K., Leung, W., He, L., Jiang, Y., Wang, C.C., Xing, G., & Chen, H. (2025). "OpenTCM: A GraphRAG-Empowered LLM-based System for Traditional Chinese Medicine Knowledge Retrieval and Diagnosis." arXiv:2504.20118.

²¹ Xiang, S. et al. (2025). "Integrating Knowledge Graphs with Ancient Chinese Medicine Classics: Challenges and Future Prospects of Multi-Agent System Convergence." *Chinese Medicine* 20, 168. <https://doi.org/10.1186/s13020-025-01226-7> -본 논문의 참고문헌 [21], [22], [23] 참조.

²² CSIR-TKDL. Traditional Knowledge Digital Library (URL: <https://www.csir.res.in/en/documents/tkdl>) ; Fredriksson, M. (2022). "India's Traditional Knowledge Digital Library and the Politics of Patent Classifications." *Third World Quarterly* 44(4), 813–832. <https://doi.org/10.1080/01436597.2022.2146529>

²³ GRAYU (URL: <https://caps.ncbs.res.in/GRAYU/>) ; Joshi, S. et al. (2025). “GRAYU: Graph-based Database Integrating Ayurvedic Formulations, Medicinal Plants, Phytochemicals and Diseases.” *Frontiers in Pharmacology* 16, 1727224. <https://doi.org/10.3389/fphar.2025.1727224>

²⁴ Polyglot Asian Medicine Knowledge Graph (URL: <https://kgraph.sg/polyglot/>) ; Khoo, C.S.G., Stanley-Baker, M., Zakaria, F.B., Chen, J., Ang, S.Q.R., & Huang, B. (2023). “Development of the Polyglot Asian Medicine Knowledge Graph System.” *ICADL 2023, LNCS 14458*, 3–11. https://doi.org/10.1007/978-981-99-8088-8_1

²⁵ Fauzi, N., Mohamed Najid, M.H., Ibrahim, N., & Sahbudin, M.A. (2025). “Digitalizing Malay Medical Manuscripts: A Database-Driven Approach to Preservation, Knowledge Management and Intellectual Property.” *Journal of Al-Tamaddun* 20(2), 199–210. <https://doi.org/10.22452/JAT.vol20no2.14>

²⁶ 본 다해석 병렬 수용 스키마(Pluralistic Interpretation Schema)의 설계 방향은 본 논문의 익명 심사 과정에서 한 심사위원이, RDM_Core 스키마에 연구자 식별자 및 해석 변인 필드 추가 등을 통한, 서로 다른 연구자의 해석적 이견이나 대안적 판단을 하나의 데이터셋 안에서 병렬로 통합·수용할 수 있는 시스템적 설계에 관한 아이디어나 방향성을 보완 서술한다면 본 연구의 데이터 모델이 갖는 확장성을 더욱 돋보이게 할 수 있을 것이라고 제시한 의견에 직접 힘입은 것임을 밝힌다. 이 제언은 본 연구가 미처 정식화하지 못한 학술적 확장의 방향을 명료하게 가리켜 주었으며, 저자는 이를 적극 수용하여 본 항목과 6.1 절 및 7 장 결론에 반영하였다.

²⁷ GRAYU (URL: <https://caps.ncbs.res.in/GRAYU/>) ; Joshi, S. et al. (2025). “GRAYU: Graph-based Database Integrating Ayurvedic Formulations, Medicinal Plants, Phytochemicals and Diseases.” *Frontiers in Pharmacology* 16, 1727224. <https://doi.org/10.3389/fphar.2025.1727224>

²⁸ FAIR Principles URL: <https://www.go-fair.org/fair-principles/>

²⁹ CARE Principles URL: <https://www.gida-global.org/care>

³⁰ 이 네 번째 학술적 기여 항목은 본 논문의 익명 심사 과정에서 한 심사위원이 제언한, 다른 연구자의 해석적 이견을 데이터셋 안에서 병렬로 통합·수용할 수 있는 시스템적 설계의 아이디어에 직접 힘입은 것이다. 본 연구는 그 제언을 수용하여 5.3 절 둘째 항목 및 6.1 절 ‘해석자 층(Interpreter Layer)’ 단락에서 그 구체적 설계 방향을 제시하였다.

참고문헌

1. 한국어 문헌

김호 (2000). <허준의 동의보감 연구>. 일지사.

류인태 (2018). “디지털 인문학과 한문학 연구 — 고문헌 자료 대상 국내 디지털 인문학 연구 사례를 중심으로”. <한문학논집>. 49. 43–76.

류인태 (2019). “데이터로 읽는 17 세기 재지사족의 일상: 『지암일기(1692~1699)』 데이터베이스 편찬 연구”. 한국학중앙연구원 한국학대학원 박사학위논문.

류인태 (2020). “디지털 인문학은 인문학이다”. <인문논총>. 77(3). 365–407.

박훈평, 안상우 (2014). “혜민서 관청지 『惠局志』 편제와 내용 연구”. <한국의사학회지>. 27(1).

신동원 (2004a). “조선시대의 의학론”. <의사학>. 13(1). 134-145.

신동원 (2004b). “조선 후기 의원의 존재 양태”. <한국과학사학회지>. 26(2).

신동원 (2015). <동의보감과 동아시아 의학사>. 들녘.

신동원, 오재근, 이기복, 전종욱 역주, 이수귀 원저 (2015). <역시만필(歷試漫筆)>. 들녘.

안상우 (2000). “의방유취에 대한 의사학적 연구”. <경희대학교 박사학위논문>.

오재근 (2022). “17-8 세기 조선의 임상 의학은 어떻게 이뤄졌을까 - 『역시만필』을 통해 살펴본 의관 이수기의 임상 의학 -”. <의사학>. 31(1). 1-34.

이경록 (2017). “『의방유취(醫方類聚)』의 편찬과 그 함의”. <국역 의방유취>. 1.

이기복 (2021). “고려·조선의 “鄉藥” 및 “東醫” 담론의 문화적 의미: 행위자의 의제 및 자아 구성 전략을 중심으로”. <한국문화>. 93. 411-454.

전종욱 (2022). <임원경제지와 조선의 일용기술>. 들녘.

전종욱 (2025). “조선에서 현대까지 한의학 교육 변천에 대한 역사적 의미 고찰과 전망: 원전 교육을 중심으로”. <대한한의학원전학회지>. 38(3). 193-213.

전종욱 (2025a). “조선의학사유구조 RDM 프로젝트 헌법”. <전북대학교 한국과학문명학연구소>.

전종욱, 조창록 (2012). “『임원경제지』·「인제지」의 편집 체제와 조선 후기 의학 지식의 수용 양상: 『동의보감』과의 비교를 중심으로”. <의사학>. 21(3). 403-448.

정명현 외 (2012). <조선 최대의 실용백과사전 임원경제지>. 씨앗을뿌리는사람.

2. 영어 문헌

ALLEA (2020). *Sustainable and FAIR Data Sharing in the Humanities*. Eds. N. Harrower et al. Berlin: All European Academies.

-
- Carroll, S.R. et al. (2020). "The CARE Principles for Indigenous Data Governance." *Data Science Journal* 19(1), p. 43. <https://doi.org/10.5334/dsj-2020-043>
- Carroll, S.R. et al. (2021). "Operationalizing the CARE and FAIR Principles for Indigenous Data Futures." *Scientific Data* 8, Article 108. <https://doi.org/10.1038/s41597-021-00892-0>
- Cheung, K.H. & Chen, H. (2010). "Semantic Web for Data Harmonization in Chinese Medicine." *Chinese Medicine* 5, p. 2. <https://doi.org/10.1186/1749-8546-5-2>
- Fauzi, N., Mohamed Najid, M.H., Ibrahim, N., & Sahbudin, M.A. (2025). "Digitalizing Malay Medical Manuscripts: A Database-Driven Approach to Preservation, Knowledge Management and Intellectual Property." *Journal of Al-Tamaddun* 20(2), 199–210. <https://doi.org/10.22452/JAT.vol20no2.14>
- Fredriksson, M. (2023). "India's Traditional Knowledge Digital Library and the Politics of Patent Classifications." *Law Critique*, 34, 1–19. <https://doi.org/10.1080/01436597.2022.2146529>
- Gordon, M. et al. (2024). "A Scoping Review of Artificial Intelligence in Medical Education: BEME Guide No. 84." *Medical Teacher* 46(4), pp. 446–470. <https://doi.org/10.1080/0142159X.2024.2314198>
- He, J., Guo, Y., Lam, L.K., Leung, W., He, L., Jiang, Y., Wang, C.C., Xing, G., & Chen, H. (2025). "OpenTCM: A GraphRAG-Empowered LLM-based System for Traditional Chinese Medicine Knowledge Retrieval and Diagnosis." arXiv:2504.20118. <https://arxiv.org/abs/2504.20118>
- Jang, H. et al. (2010). "Ontology for Medicinal Materials Based on Traditional Korean Medicine." *Bioinformatics* 26(18), pp. 2359–2360. <https://doi.org/10.1093/bioinformatics/btq424>
- Joshi, S. et al. (2025). "GRAYU: Graph-based Database Integrating Ayurvedic Formulations, Medicinal Plants, Phytochemicals and Diseases." *Frontiers in Pharmacology* 16, 1727224. <https://doi.org/10.3389/fphar.2025.1727224>
- Kerry, C.F. et al. (2026). *Is AI Sovereignty Possible? Balancing Autonomy and Interdependence*. Brookings Institution Report.
- Khoo, C.S.G., Stanley-Baker, M., Zakaria, F.B., Chen, J., Ang, S.Q.R., & Huang, B. (2023). "Development of the Polyglot Asian Medicine Knowledge Graph System." *ICADL 2023, LNCS 14458*, 3–11. https://doi.org/10.1007/978-981-99-8088-8_1
- Kim, S.K. et al. (2015). "TM-MC: A Database of Medicinal Materials and Chemical Compounds in Northeast Asian Traditional Medicine." *BMC Complementary and Alternative Medicine* 15, Article 218. <https://doi.org/10.1186/s12906-015-0758-5>

Liu, C. et al. (2024). "Research on Traditional Chinese Medicine: Domain Knowledge Graph Completion and Quality Evaluation." *JMIR Medical Informatics* 12, e55090. <https://doi.org/10.2196/55090>

Long, H. et al. (2019). "An Ontological Framework for the Formalization, Organization and Usage of TCM-Knowledge." *BMC Medical Informatics and Decision Making* 19(Suppl. 2), p. 53. <https://doi.org/10.1186/s12911-019-0760-9>

Shin, D. (2010). "How Commoners Became Consumers of Naturalistic Medicine in Korea, 1600–1800." *East Asian Science, Technology and Society* 4(2), pp. 275–301. <https://doi.org/10.1215/s12280-010-9135-0>

Singh, S.K. et al. (2025). "Sovereign AI: Rethinking Autonomy in the Age of Global Interdependence." arXiv:2511.15734. <https://arxiv.org/abs/2511.15734>

Wang, S. et al. (2024). "A Review of Knowledge Graph in Traditional Chinese Medicine." *Computers, Materials & Continua* 81(3). <https://www.techscience.com/cmc/v81n3/59025/html>

Wilkinson, M.D. et al. (2016). "The FAIR Guiding Principles for Scientific Data Management and Stewardship." *Scientific Data* 3, Article 160018. <https://doi.org/10.1038/sdata.2016.18>

Xiang, S. et al. (2025). "Integrating Knowledge Graphs with Ancient Chinese Medicine Classics: Challenges and Future Prospects of Multi-Agent System Convergence." *Chinese Medicine* 20, 168. <https://doi.org/10.1186/s13020-025-01226-7>

Zhou, X. et al. (2004). "Ontology Development for Unified Traditional Chinese Medical Language System." *Artificial Intelligence in Medicine* 32(1), pp. 15–27. <https://doi.org/10.1016/j.artmed.2004.01.014>

3. 온라인 리소스

전종욱 (2025b–2026). 「조선의학 질병 인식구조 RDM — 30 병문」. KISTI DataON 국가연구데이터플랫폼. <https://dataon.kisti.re.kr>

〈전통의학정보포털〉 URL: <https://oasis.kiom.re.kr>

〈한국전통지식포털〉 URL: <https://www.koreantk.com>

〈한의학융합연구정보센터〉 URL: <https://www.kmcric.com>

〈한의학고전 DB〉 URL: <https://mediclassics.kr>

〈CARE Principles〉 URL: <https://www.gida-global.org/care>

〈CSIR-TKDL〉 URL: <https://www.csir.res.in/en/documents/tkdl>

〈FAIR Principles〉 URL: <https://www.go-fair.org/fair-principles>

〈GRAYU〉 URL: <https://caps.ncbs.res.in/GRAYU/>

〈Polyglot Asian Medicine Knowledge Graph〉 URL: <https://kgraph.sg/polyglot/>

부록 A. 30 병문 데이터셋 일람표

본 부록은 본 연구의 30병문 데이터셋이 KISTI DataON 국가연구데이터플랫폼에 등록된 현황을 한눈에 정리한 것이다. 각 병문은 개별 DOI와 함께 등록·공개되어 있어 직접 접근이 가능하다. DataON 홈페이지(<https://dataon.kisti.re.kr>)에서 ‘조선의학 질병 인식 구조 RDM’ 또는 개별 병문명을 검색하면 즉시 도달할 수 있다. (표기: ★ = STAGE3 시트 구축 / ◇ = 『역시만필』 의안 또는 임상 사례 직접 반영.)

계열	병문 (30)
외감·시령 (7)	해수문 ◇, 해학문, 온역문, 습병문, 중서문, 중한고행문, 조열문
체액·순환 (5)	부종문, 담음문, 혈병문 ◇, 제기문 ◇, 적취문 ◇
소화·대사 (7)	곽란문 ◇, 구토문 ◇, 창만문 ◇, 황달문 ★, 소갈문 ★, 대변문 ◇, 내상문
허손·정신 (2)	허로문 ◇, 경계정충견망문
감각·배설 경계 (6)	안문, 이문, 비문, 치문, 소변문, 전간광문
외상·구급 (3)	해독문, 금창문, 제상문

『역시만필』 +1 적용은 30병문 중 11병문(약 37%)에 이루어졌으며, 그 중 소갈문·황달문은 STAGE3 시트로 별도 구축되어 임상 교차 검증이 가장 두텁게 이루어졌다. 본 30병문은 본 연구의 중착점이 아니라 전체 병문(50병문 정도로 예상됨)으로 나아가는 과정의 단계적 한 단면이다.

부록 B. 조선의학 사유구조 RDM 프로젝트 헌법

본 연구의 데이터셋 구축 원칙과 운영 규약을 정리한 「조선의학 사유구조 RDM 프로젝트 헌법」을 본 논문의 별첨 문서로 제공한다. 본 헌법은 (1) 프로젝트 목적, (2) 3축 + 1 문헌 체계, (3) 기본 원칙(문헌 우선·분리·구조화), (4) 분석 단위(병문), (5) L1·L2·L3 판단 층위, (6) Type A-D 분류, (7) Provenance_Citation 원칙, (8) RDM 데이터 스키마, (9) 데이터 패키지 구성(All-in-One), (10) 번역 원칙, (11) NETWORK 관계어휘 정책, (12) 본 헌법의 범위(30병문 목록), (13) DataON 등재 정책, (14) 프로젝트 비전, (15) 다 해석 병렬 수용을 위한 후속 확장 비전의 15장으로 구성된다.