

石泉 申綽의 『尙書古注』 연구 -경학자료 DB 구축을 위한 시론-

이동학¹⁾

1. 서론
2. 『尙書古注』 온톨로지 디자인
3. Graph DB로 『尙書古注』 읽기
4. 맺음말

1. 서론

본 연구는 石泉 申綽(1760~1828)의 『尙書古注』를 대상으로 시맨틱 데이터 프로세싱을 수행함으로써, 경학 자료에 내재한 지식 관계망을 구조화하고 해석 논리와 주석 간 상호작용을 분석할 수 있는 데이터 모델을 제시하는 데 목적이 있다.²⁾

『상서고주』는 『今文尙書』를 대상으로 漢~唐代 경학가들의 학설을 집록한 저술이다. 신작은 『금문상서』만을 正經으로 인정하였기 때문에 『금문상서』를 해설하기 위한 주석서를 따로 저술하였다.³⁾ 주로 언급되는 학자로는 歐陽, 夏侯, 馬融, 鄭玄, 王肅 등이 있고, 『尙書大傳』, 『尙書正義』, 『史記』, 『漢書』, 『後漢書』, 『說文解字』, 『三國志』, 『文選註』 등의 서적을 인용하였다. 신작은 본인의 견해를 직접 개진하기보다 여러 학자들의 설과 문헌의 내용을 언급해 두는 방식을 채택하였다.

경학 자료는 경문과 주석 간의 층위가 존재하고, 주석 간에도 상호 맥락이 형성되어 있다. 경문에 대한 주석은 字句를 풀이하는 訓詁, 우주론·심성론으로 개념을 확장시켜 해석하는 義理, 경문을 여타 문헌을 통해 실증적으로 입증하고자 하는 考證의 성격을 지닌다. 이처럼 경학 자료는 다층적·다면적 구조를 지니고 있다는 점에서, 온톨로지를 기반으로 한 설계를 통해 디지털 환경에서 구현하기 적합한 특성을 지닌다. 개별 주석의 관계를 설정하고 그것이 경문과 어떤 양상으로 연결되는지를 고찰하는 것은 단지 경문의 이해에 기여할 뿐 아니라 주석가가 경전을 이해하는 총체적인 사고구조를 드러내는 데에도 중요한 의미를 지닌다. 이는 아날로그적 연구 환경에서는 체계적으로 조직하기 어려웠던 경학 자료의 다층적 관계망을 디지털 환경에서 유연하고 풍부하게 재구성할 수 있는 가능성을 열어준다.

1) 고려대학교 국어국문학과 한문학전공 박사수료.

2) 『尙書古注』의 XML 스키마 설계는 변은미·이동학·유인태(2024), 『『尙書古訓』과 『尙書古注』 연계 시맨틱 데이터 프로세싱 1 : XML 데이터 설계 및 편찬』, 『한문학논집』 69, 근역한문학회 참고.

3) 신작의 저술로는 『詩經』 관련 저술인 『詩次故』, 『周易』 관련 저술인 『易次故』가 있고 『尙書』 관련 저술로는 본고에서 다룬 『상서고주』 외에 『古尙書』, 『尙書二十五篇』, 『尙書百篇攷』가 있다. 『고상서』는 『금문상서』 28편의 본문을 隸書體로 기록하여 『금문상서』의 원형을 복원하고자 하였다. 『상서이십오편』은 僞『古文尙書』 25편의 원문과 주석을 모은 것으로 僞書라든 문장이 매우 전 아하여 경학연구에 도움이 되는 것으로 보았다. 『상서백편고』는 『상서』 100편의 『書序』에 대한 고증을 한 것으로, 『상서정의』에 나오는 「서서」를 저본으로 『사기』, 『한서』 등에 나오는 「서서」와 기타 경전에서 나타난 구절들을 덧붙여 정리하였다.

특히 『상서』는 『금문상서』와 『고문상서』 논쟁, 다양한 異文의 발생, 逸失된 경문의 복원 시도 등 경학사적 쟁점이 각 자구마다 함축적으로 담겨 있는 텍스트이다. 이러한 점에서 연구자의 해석적 관점을 바탕으로 온톨로지를 설계한다면, 『상서』 학술사의 전반적인 흐름과 경전을 복원하고 해석하고자 하는 주석가의 악전고투를 디지털 환경 속에서 구현할 수 있을 것으로 기대된다.

이에 2장에서는 『상서고주』의 온톨로지 전반과 설계 지침에 관한 내용을 다루고, 그것이 실제 시각화된 그래프로 구현된 양상을 3장에서 보이고자 한다.⁴⁾ 이를 통해 하나의 경학 자료가 디지털 환경에서 구현된 예시를 보이고, 향후 경학 데이터베이스가 나아갈 방향을 논의해보고자 한다.

4) 『尙書古注』 「堯典」 1구~69구를 대상으로 하였다.

2. 『尙書古注』 온톨로지 디자인

온톨로지(Ontology)는 본래 철학의 ‘존재론’으로, 존재의 본질과 그 구조를 탐구하는 학문이다. 정보과학에서는 이 개념이 확장되어 ‘정보화의 대상이 되는 세계를 전자적으로 표현할 수 있도록 구성한 데이터 기술 체계’⁵⁾를 의미한다. 즉 디지털 환경에서 개별 요소와 요소들 간의 관계, 그리고 이들의 속성이 유기적으로 결합된 데이터의 총체로 이해할 수 있다.

디지털 인문학의 영역에서 온톨로지에 기반한 연구는 지속적으로 이루어져 왔으며, 특히 한문학 분야에서는 장르별로 다양한 연구 성과가 제출되었다.⁶⁾ 인문 데이터가 지닌 연속적이면서도 다층적인 구조를 디지털 환경에서 효과적으로 구현하기 위한 과정으로서 온톨로지의 설계가 유효한 방법론이었음을 확인할 수 있다.

다음은 『상서고주』의 클래스(Class), 속성(Attribute) 및 관계(Relation)를 설계한 것이다.

2-1. 노드(Node) 설계

1) 원문 클래스

Class	Attribute	예시
Section	id	S1
	name	우하서
	content	虞夏書
Volume	id	V1
	name	권1
	content	卷一
Chapter	id	C01
	name	요전
	content	堯典
Line	id	C01L001
	name	요전1구
	content	曰若稽古

『상서고주』는 서(書: 虞夏書、商書、周書), 권(卷: 卷1~卷8), 편(篇: 堯典、大禹謨 등), 구(句)로 구성된다.

5) 김현·안승준·유인태(2018), p. 23.

6) 대표적 시맨틱 데이터 모델링 사례로는 누정기(박순, 2017), 고문서(김현·안승준·유인태, 2018), 생활일기(유인태, 2019), 야담(양승목·유인태, 2022), 여향문학(김지선·유인태, 2022), 시화(한희연·지영원·호정·유인태, 2022), 한시(지영원·최진경, 2024), 소설(이길환, 2024), 경학(변은미·이동학, 2024), 과시자료(이길환·변은미·유인태, 2024), 관원록(김바로, 2023), 그 외 한국문학(장문석·유인태, 2021), 불교학(김바로, 2018) 등을 들 수 있다.

2) 인용(Citation) 클래스

Attribute	구분	설명	예시
id	식별자	원문 id+언더바+영역(G/C)+CI+순서	C01L007_CCI4
name	대표명	원문 name+언더바+영역명+순서+인용	요전7구_교감4인용
area	영역	훈고·교감·분류표지	교감
form	형태	기본·소주·난외주	기본
format	형식	직접인용·간접인용	직접인용
context	맥락	단순인용·긍정·부정·부분긍정 참조·근거·비판	단순인용
content	내용	인용 부분 전체 입력	徐廣曰 駟古訓字 順也

3) 견해(Opinion) 클래스

Attribute	구분	설명	예시
id	식별자	원문 id+‘_’+영역+O+순서	C01L017_CO1
name	대표명	원문 name+‘_’+영역+순서+견해	요전17구_교감1견해
content	내용	견해 부분 전체 입력	綽按 …

신작은 ‘綽按’의 표지를 사용하여 견해를 제시하였으나 대부분 인용 주석의 나열로 구성되어 있다. 견해를 명시적으로 개진하기보다는 기존의 다양한 인용을 통해 암묵적이고 우회적으로 자신의 입장을 드러내는 방식을 취하였다.

4) 언급(Mention) 클래스

Attribute	구분	설명	『상서고주』 예시
id	식별자	인용 id+MT+순서	C01L069_CMT1
name	대표명	인용 name+순서+언급	요전69구_교감1언급
area	영역	훈고·교감·분류표지	교감
content	내용	언급 부분 전체 입력	今文及馬鄭王本 ……爲舜典之初

언급 클래스는 인용 중간중간에 간헐적으로 보이는 신작의 언급을 담는 틀이다.

5) 구절(Text) 클래스

Attribute	구분	설명	『상서고주』 예시
id	식별자	인용 id+T	C01L005_CQ1T
name	대표명	인용 name+구절	요전5구_교감1인용구절
content	내용	구절 부분 전체 입력	合於堯之克攘
reference	출처	구절과 함께 제시된 출처를 입력	漢藝文志
reference type	출처 유형	제시된 출처의 유형을 입력	Book, Part, Person
reference target	출처 대상	제시된 출처의 온전한 명칭을 입력	漢書 藝文志

인용된 정보의 출처를 인물, 문헌, 편명 등 다양한 수준에서 정밀하게 식별할 수 있도록 하여 출처 기반의 의미적 연결 관계를 구성하는 기반이 된다.

6) 문헌(Book) 클래스

Attribute	구분	설명	예시
id	식별자	B+구별 번호(입력순)	B001
korname	한글명	서명의 한글명 입력	서찬
chiname	한자명	서명의 한자명 입력	書贊

문헌 클래스는 문헌 정보를 담는 틀로, 인용 출처에 해당하는 문헌, 연구 대상으로 삼은 경서 주석서, 단순 언급된 문헌까지 모두 포함된다.

7) 인물(Person) 클래스

Attribute	구분	설명	예시
id	식별자	P+구별 번호(입력순)	P001
korname	한글명	인물의 한글명 입력	정현
chiname	한자명	인물의 한자명 입력	鄭玄
birthyear	생년	인물의 생년을 입력	127
deathyear	몰년	인물의 몰년을 입력	200
nationality	국적	한국 · 중국 · 일본 구분	중국

8) 금고문(NewOld) 클래스

Attribute	구분	설명	예시
id	식별자	NO+순서	NO008
name	대표명	금고문+순서	금고문008
content	내용	금고문 부분 전체 입력	今文及馬鄭王本 堯舜典合爲一 梅賾所上孔傳古文 猶闕舜典 乃分慎徽五典以下 爲舜典之初

금고문 클래스는 고문상서·금문상서 관련 내용을 담는 틀이다.

9) 비문(碑文, Epigraph) 클래스

Attribute	구분	설명	예시
id	식별자	Ep+순서	Ep005
korname	한글명	비(碑)의 한글명 입력	당부비
chiname	한자명	비(碑)의 한자명 입력	唐扶碑
content	내용	비문의 내용 입력	追惟堯德 廣被之恩

2-2. 관계(Relation) 설계

관계 데이터는 경학 자료를 구성하는 정보들을 의미적으로 연결하기 위한 것으로 원문과 주석, 주석에 등장하는 인용문과 출처, 원문 및 인용문에 관한 견해 등에 관한 관계를 제시하는 것이다. 이는 경학자가 경문을 설명하기 위해 사용한 인용문의 맥락, 자료 배치 의도, 특정 쟁점에 관한 경학자의 논증 및 사고 과정을 추적하는 데 도움을 준다.

유형	Source(A)	Target(B)	Relation	설명
원문-원문	원문	원문	hasPart	A는 B를 부분으로 가지다
주석-원문	인용	원문	is Cited as reference for	단순 인용
	견해	원문	is Opinion of	A는 B에 대한 견해이다
주석-주석 (기본형)	인용	인용	supplement	A는 B를 보충하다
			contradict	A는 B를 반박하다
			is Aligned with	A는 B와 같은 입장이다
주석-주석 (내포형)	인용	인용	is Cited by/ is Cited to support	참조 인용/ 근거 인용
	견해	인용		
주석-주석 (종속형)	언급	인용	is Mention for	A는 B에 대한 언급이다
인용 내	구절	인용	is Text of	A는 B에 대한 구절이다
	문헌	구절	mention	A는 B를 말하다
	인물	구절		
	인물	문헌	is Writer of	A는 B의 저자이다
	문헌	인물	contain	A는 B를 포함하다
주석-해석	인용	금고문	include	A는 B를 포함하다

1) 원문-원문 관계

자료별 분절 단위에 따라 클래스를 생성하고, 각 단위는 **hasPart(포함하다)**의 관계어를 통해 계층적으로 연결하였다. 『상서고주』는 ‘편’ 하위에 ‘장’이 아닌 ‘구’가 바로 귀속된다.

2) 주석-원문 관계

원문에 대한 주석의 형태가 특정 문헌의 문장 혹은 인물의 언급을 인용하였을 경우에는 is Cited as reference for(레퍼런스를 인용하다), ‘緯按’이라는 표지가 있어서 신작의 견해가 개입된 것을 확인할 수 있는 경우에는 is Opinion of(견해이다)의 관계어를 설정하였다.

3) 주석-주석 관계(기본형)

■ 보충 관계 (supplement)

- 요전5구_교감1인용 : 漢藝文志, 合於堯之克攘.
- 요전5구_교감2인용 : 注師古曰, 攘古讓字.

■ 반박 관계 (contradict)

- 요전8구_훈고2인용 : 許慎五經異義, …… 明在九族中, 不得但施於同姓.
- 요전8구_훈고3인용 : 鄭玄駁云, 異姓之服, 不過緦麻. 言不廢昏. 又昏禮請期云, 惟是三族之不虞. 恐其廢昏, 明非外族也.

■ 동일 관계 (is Aligned with)

- 요전3구_훈고1인용 : 馬融曰, 放勳堯名.
- 요전3구_훈고2인용 : 趙歧孟子注, 同.

4) 주석-주석 관계(내포형)

■ 근거 (is cited to support)

- 요전43구_교감1견해 : 綽按, 庸, 用也, 用可乎, 言不可用也. 史記, 頑凶不用, 卽馬義也.
- 요전43구_교감1견해1인용 : 史記, 頑凶不用.

■ 참고 (is cited by)

- 요전11구_교감3인용 : 師古曰, 今尙書作變, 此紀作蕃, 兩說並通.
- 요전11구_교감4인용 : 今尙書作變.

5) 인용 내 관계

구절-인용 간의 연결은 is Text of(구절이다)의 관계어로 연결하였다. 문헌-구절, 인물-구절 간의 연결은 모두 mention(말하다)의 관계어로 설정하였다. 저자와 문헌 간의 관계는 is Writer of(저자이다), 문헌과 발화자 간의 관계는 contain(포함하다)로 연결하였다.

6) 주석-해석 관계

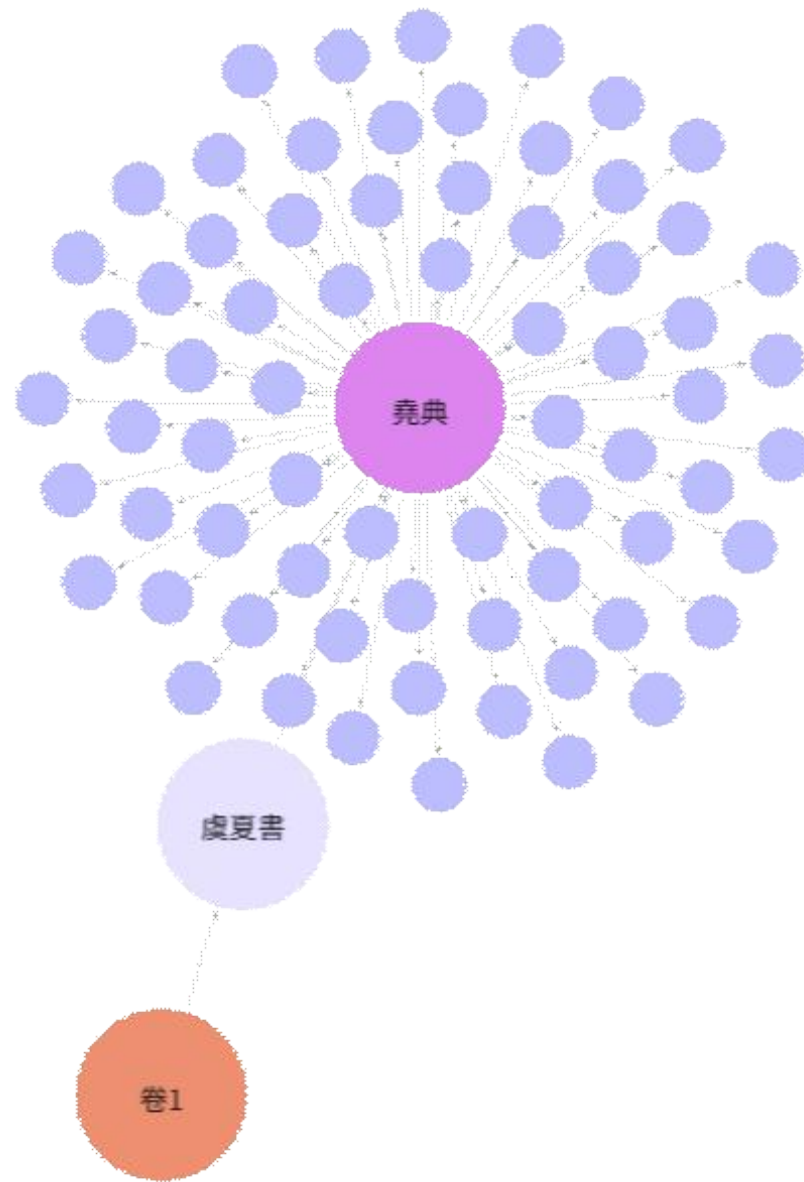
금고문을 포함하고 있는 인용을 연결하여 금고문 관련 담론을 파악할 수 있도록 하였다.

3. Graph DB로 『尙書古注』 읽기

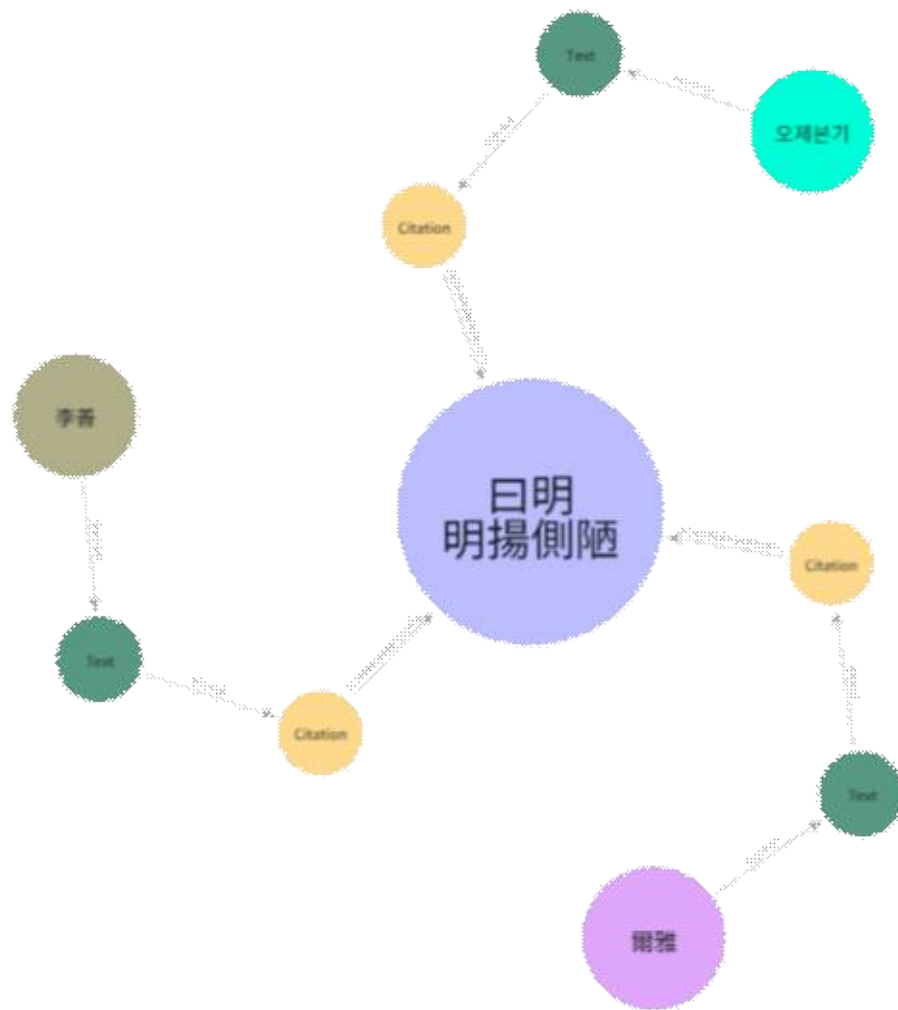
3-1. Neo4j로 구현한 『尙書古注』



[그림 1] 모든 Node와 관계망



[그림 2] Volume-Section-Chapter-Line



[그림 3] Line-Citation-Text-Book · Person · Part

3-2. Query를 통한 데이터 시각화 및 추출

- * Query를 입력하여 직접 시연하겠습니다.
- * Graph DB는 Neo4j Aura를 사용하였습니다.
- * Query 작성은 ChatGPT의 도움을 받았음을 밝힙니다.

- 전체 Node와 관계 조회:

```
MATCH p=()-[]->() RETURN p
```

- 각 구(句)별 인용 수 및 양상:

```
MATCH p=(a: Citation)-[r: isCitedAsReferenceFor]->(b: Line) RETURN a, r,
```

- 각 구(句)별 인용 수 및 양상(1순위~10순위)

```
MATCH (a: Citation)-[:isCitedAsReferenceFor]->(b: Line)
RETURN
b.gid, b.content      AS 경문,
COUNT(a)            AS 주해건수
ORDER BY 주해건수 DESC
LIMIT 10;
```

- 정현과 허신 사이 3단계 이내로 연결되는 Book, Part, Person

```
MATCH path = (p:Person)-[*..3]-(related)
WHERE p.gid IN ['P001', 'P006']
RETURN DISTINCT path
```

- 훈고가 인용된 Line

```
MATCH (a:Citation {area:"Glossary"})-[r:isCitedAsReferenceFor]->(b:Line)
RETURN a, r, b
```

- 교감이 인용된 Line(異文의 발생)

```
MATCH (a:Citation {area:"Collation"})-[r:isCitedAsReferenceFor]->(b:Line)
RETURN a, r, b
```

- 금문·고문 관련 내용이 언급된 주석과 구

```
MATCH (l:Line)<-[:isCitedAsReferenceFor]-(a:Citation)-[:include]->(b:NewOld)
RETURN l, r1, a, r2, b
```

- 비문(碑文)이 언급된 주석과 구

```
MATCH (a: Epigraph) - [r1:mention] -> (b:Text) - [r2:isTextof] -> (c:Citation) -
[r3:isCitedAsReferenceFor] -> (d: Line) RETURN a, r1, b, r2, c, r3, d
```

- 언급과 금고문

```
MATCH p=()-[:isMentionOf]->()-[:]->() RETURN p
```

- 견해(綽按)과 언급

```
MATCH p=()-[:isMentionOf]->(Citation {area:"Opinion"})-[:]->() RETURN p
```

- Book, Part, Person에서 가져온 인용문과 관계

```
MATCH p=()-[:mention]->() RETURN p
```

- 본문과 주석이 함께 인용

```
MATCH p=()-[:Supplement]->()-[:]->() RETURN p
```

- Citation간 관계 존재

```
MATCH      p=(a:Citation)-[r]->(b:Citation)-[r2:isCitedAsReferenceFor]->(c:Line)
RETURN a, r, b, r2, c
```

4. 맺음말

이상으로 『상서고주』의 온톨로지를 설계하고 「요전」의 일부를 대상으로 그래프 DB를 구현하였다. 그래프DB의 장점은 경전 주석서를 구성하는 각 요소들과 각각의 연결 관계를 통합적으로 조망하거나 연구자의 필요에 따라 일부 Node와 관계망을 집중적으로 살펴볼 수 있다는 점이라 하겠다.

그러나 본 연구는 아직 해결하지 못한 문제점들을 상당히 내포하고 있다. 우선 그래프 DB를 구성하는 과정에서 데이터가 정제되지 않았고 이로 인해 일부 relation 혹은 id 등에서 오류가 발생하였다. 이는 추후 바로잡아야 할 것이다.

또한 경학자료를 데이터로 입력하는 목적은 아날로그적 연구환경에서 발견하기 힘든 자료의 학술사적 가치를 디지털 도구를 이용하여 추출하고 구현해보고자 함이다. 경학 자료의 디지털화가 유효한 방법론으로 자리잡기 위해서는 ‘어떤 Query를 만들어야 가치있는 데이터를 추출해낼 수 있는가?’에 대한 고민이 선행되어야 하겠다.

마지막으로 Graph DB에서 원하는 데이터를 추출해내기 위해서는 연구자가 모든 Node와 Relation, 속성값을 모두 알고있는 상태에서 이를 이용한 Query를 작성하여야 한다. 즉 ‘연구를 설계한 연구자가 검색하기 편한 시스템’이다.

그러나 Query의 작성은 복잡하고 번거로우며 ChatGPT 등의 LLM의 도움을 받는다 하더라도 오류가 잦은 문제가 있다. 이에 한국고전종합DB, 한국경학자료시스템 등과 변별되는 DB로서의 의미를 갖추기 위해서는 보다 범용성 있는 데이터 탐색 방안이 강구되어야 할 것이다.

◎ 참고문헌

<학술논문>

- 김현·안승준·류인태(2018), 「데이터 기반 인문학 연구 방법의 모색: 문중 고문서 아카이브와 디지털 인문학의 만남」, 『횡단인문학』 창간호, 숙명여자대학교 숙명인문학연구소.
- 김바로(2018), 「디지털 인문학의 관점에서 본 디지털 불교학」, 『불교학보』 86, 동국대학교 불교문화연구원.
- 유인태(2019), 「데이터로 읽는 17세기 재지사족의 일상-『지암일기(1692-1699)』 데이터베이스 편찬 연구」, 한국학중앙연구원 박사학위논문.
- 장문석·유인태(2021), 「디지털 인문학과 한국문학 연구(1) : 작가 연구를 위한 시맨틱 데이터베이스 설계」, 『민족문학사연구』 75, 민족문학사학회.
- 김지선·유인태(2022), 「여향문화 연구와 데이터 모델링」, 『한국한문학연구』 85, 한국한문학회.
- 양승목·유인태(2022), 「야담의 데이터, 야담으로부터의 데이터: 한국 야담 데이터 모델의 구상」, 『한국문학연구』 68, 동국대학교 한국문학연구소.
- 한희연·지영원·호정·류인태(2022), 「한국 詩話 시맨틱 데이터 아카이브 구현 방안 모색 -『詩話叢林』 대상 개념적 데이터 모델링을 중심으로」, 『한문학논집』 63, 근역한문학회.
- 김바로(2023), 「한국역사인물데이터베이스 설계 시론: 홍문록(弘文錄)을 예시로 하여」, 『한국학』 46, 한국학중앙연구원.
- 변은미·이동학·유인태(2024), 「『尙書古訓』과 『尙書古注』 연계 시맨틱 데이터 프로세싱 1 : XML 데이터 설계 및 편찬」, 『한문학논집』 69, 근역한문학회.
- 이길환(2024), 「한국한문소설 시맨틱 데이터 모델링에 대한 연구 : 한국한문소설 DB 구축을 위한 예비적 검토」, 『민족문학사연구』 85, 민족문학사연구소.
- 이길환·변은미·유인태(2024), 「조선시대 科試 자료와 시맨틱 데이터 모델링」, 『한국한문학연구』 92, 한국한문학회.

<웹 자료>

- COD-Project Repository URL: <https://github.com/KU-HJHM/COD-Project>
- 한국고전종합DB URL: <https://db.itkc.or.kr/>
- 한국경학자료시스템 URL: <http://koco.skku.edu/>
- 상서고주 노드 데이터 URL:
<https://docs.google.com/spreadsheets/d/1YuYx0oaiOgcTWT6otwSwk9wwl512gWSutWm3U7BMDCQ/edit?usp=sharing>
- 상서고주 옛지 데이터 URL:
https://docs.google.com/spreadsheets/d/1hBZV4tmfzOuZEHRmhtvv1NV3HI5w_dydBLX-mRGDDVv/edit?usp=sharing