

# 전자자원의 BIBFRAME 전환 구축


발제: 문헌정보학전공 박사수료 박미진


Articles

## Research on Knowledge Organization of Bibliographic Data of Library's Electronic Resources Based on BIBFRAME

Jinghua Cui   & Mark Xuefang Zhu 

Pages 107-119 | Published online: 24 Nov 2025

 Cite this article

 <https://doi.org/10.1080/00987913.2025.2581427>



### Abstract

This study constructs a bibliographic data ontology of library's electronic resources, reusing the BIBFRAME ontology. Firstly, the mapping of classes and attributes between bibliographic data and BIBFRAME fields is investigated. And conceptual models are drawn from the work layer, instance layer, item layer and connection fields. Then the constructed bibliographic data ontology of electronic resources has been visualized by using Protégé, and Neo4j has been used to respectively implement the storage and querying of publications, thesis and other electronic resources, such as patents.

### Keywords:

Bibliographic data

library's electronic resources

BIBFRAME

knowledge organization

본 연구는 BIBFRAME 온톨로지를 재사용하여 도서관 전자자원의 서지데이터 온톨로지를 구축한다. 먼저, 서지데이터와 BIBFRAME 필드 간의 클래스 및 속성 매핑을 조사하였다. 또한 저작 계층, 인스턴스 계층, 아이템 계층 및 관계 필드에서 개념 모델을 도출하였다. 그 후 구축된 전자 자료 서지 데이터 온톨로지를 Protégé를 사용하여 시각화하였으며, Neo4j를 활용하여 간행물, 학위 논문 및 특허와 같은 기타 전자 자료의 저장 및 쿼리 기능을 각각 구현하였다.

# 연구의 배경

- 스마트도서관(smart library)으로의 전환
  - 개인화되고, 능동적이며, 정확하고 포괄적인 지식서비스 제공(이용자 & 사서)
- 서지데이터의 지식조직(knowledge organization)의 중요성
  - 특히 전자자원 서지데이터의 조직의 필요성
    - 이용자들의 독서 습관의 변화와 전자독서(e-reading) 선택 이용자의 증가
- MARC의 한계와 BIBFRAME의 등장과 발전
  - 서지데이터 간의 의미적 관계(semantic relationships)를 보다 효과적으로 표현
- BIBFRAME 데이터 변환 과정에서 나타나는 주요 과제(PCC 보고)
  - 원시 데이터(original data)를 생성할 때 BIBFRAME 온톨로지(ontology)를 표현하는 방식의 차이
  - MARC 데이터 변환 결과의 차이

# 연구설계

- 연구목적
  - 서지데이터의 BIBFRAME 변환 과정에서 데이터의 출처 규명
  - CNMARC에서 BIBFRAME으로 변환하는 과정에서 발생하는 문제점 탐색 및 해결방안 탐색
- 선행연구
  - MARC 데이터와 BIBFRAME 간의 개체 및 관계 매핑
    - 인쇄자료의 서지데이터에 초점을 맞추고 있어 전자자료의 서지데이터에 대한 연구 부족
  - 도서관 전자자료의 활용, 영향, 성과평가, 서비스 관련 연구
    - 목록 관련 연구는 주로 CNMARC 서지 포맷에 기반한 목록 연구에 머무른 수준
  - 인공지능(AI) 기반 서비스를 제공하기 위한 지능형 시스템(intelligent systems) 개발
    - 전자자료 서지데이터의 지식조직에 관한 연구 수행의 필요성 강화
  - 도서관 분야에서의 BIBFRAME 적용 연구
    - MARC, RDA, BIBFRAME 간의 매핑과 변환
- 연구방법: 탐색적 연구
  - 전자자료의 MARC 기반 서지데이터와 BIBFRAME 온톨로지 간의 클래스(classes) 및 속성(attributes) 매핑
    - 개념 정의: 전자자료이란 ‘컴퓨터와 네트워크의 도움을 받아 온라인에서 이용하거나 다운로드해야 하는 자원(SuYong, 2023)’
    - 연구데이터: 난징도서관 구독 DB CNKI에서 전자자료의 서지데이터 수집(출판물, 논문, 기타자료)

# 연구내용 ① 전자자원 서지데이터 온톨로지의 클래스와 속성 매핑

## • 전자자원의 저작 유형과 매핑 관계 정의

### • bf:Work의 하위 클래스 설정

- **Publication**: 연속간행 학술지, 학술대회 프로시딩, 연속 신문간행물, 연감 간행물
- **Thesis**: 학술논문·학위논문·학술대회논문
- **Other electronic resources**: 도서, 신문, 연감, 연구성과, 특허, 표준

### 서지데이터-BIBFRAME 매핑 유형

- 일대일(one-to-one)
- 일대다(one-to-many)
- 다대일(many-to-one)

### 서지데이터-BIBFRAME 매핑 과정

- 서지데이터는 BIBFRAME의 클래스에만 아니라, 속성에도 매핑될 수 있음
  - 예) 학술대회 개최일(conference time)

**Table 1. Mapping of e-resource work categories.**

Work	Type	Work	Type
Serial Journal Publications	bf: Series	Journal Article	bf: Text
Conference Proceedings	bf: Collection	Conference Paper	bf: Text
Serial Newspaper Publications	bf: Series	Newspapers	bf: Text
Almanac Publications	bf: Integrating	Almanacs	bf: Text
Dissertations	bf: Text	Books	bf: Monograph
Patents	bf: Text	Standards	bf: Text
Outcomes	bf: Text		

### 클래스 및 속성의 적용

- bf:Work, bf:Instance, bf:Item 모두에서 동시에 사용 가능한 클래스 및 속성도 존재

## • bf:Instance의 하위 클래스 bf:Electronic으로 전자자원의 인스턴스 매핑

# Publication의 매핑

- 출판물(publications)의 BIBFRAME 클래스와 속성 매핑
  - 연속간행 학술지(serial journal publications), 학술대회 프로시딩(conference proceedings), 연속신문간행물(serial newspaper publications), 연감 간행물(almanac publications)

**Table 2.** Mapping of publication classes and attributes.

Work Level	bf: Work	Work Level	bf: Work
Title, Section name, Column name	bf: Title	Previous titles of journals and yearbooks	bf: variantType
Name of organizer, Supervisor, Editor-in-chief	bf: Contribution	Language type	bf: Language
Subject term	bf: Topic	Classification number	bf: Classification
Description of contents	bf: Summary	Title of the meeting	bf: Event
Meeting time	bf: date	Venue of the conference	bf: Place
Year of yearbook	bf: Temporal	Catalogs	bf: TableOfContents
<b>Instance Level</b>	<b>bf: Instance</b>	<b>Instance Level</b>	<b>bf: Instance</b>
Publication information, Publisher, Place of publication	bf: ProvisionActivity	Publication period	bf: Frequency
Pagination	bf: Extent	Identifiers	bf: Identifier
Book format	bf: BookFormat	Date of publication, Date of founding of the journal	bf: date
Statement of responsibilities	bf: responsibilityStatement	Paperback pricing	bf: acquisitionTerms
<b>Item Level</b>	<b>bf: Item</b>	<b>Item Level</b>	<b>bf: Item</b>
Times	bf: Chronology	Enumeration	bf: Enumeration

### 클래스 및 속성의 확장 및 적용

- bf:Classification > CLO
- bf:Identifier > CN, 우편번호(postal code), 판차(edition number)
- 전자출판물의 시간적 순차표시(temporal enumeration) 방식에 따라 제시하기 위한 적용
  - bf:Chronology
  - bf:Enumeration

# Thesis의 매핑

- 논문(thesis)의 BIBFRAME 클래스와 속성 매핑
  - 학술지 논문(journal articles), 학위논문(dissertations), 학술대회논문(conference papers)

**Table 3.** Mapping of Thesis classes and attributes.

Work Level	bf: Work	Work Level	bf: Work
Title	bf: Title	Language type	bf: Language
Author	bf: Contribution	Author affiliations	bf: Organization
Academic advisor	bf: Role	Degree information	bf: Dissertation
Degree	bf: degree	Graduating institution	bf: grantingInstitution
Degree year	bf: validDate	Academic program	bf: subject
Title of the meeting	bf: Event	Meeting time	bf: date
Venue of the Conference	bf: Place	Summary	bf: Summary
Keywords	bf: Topic	Classification number	bf: Classification
Funding	bf: AcquisitionSource	Article catalog	bf: TableOfContents
<b>Instance Level</b>	<b>bf: Instance</b>	<b>Instance Level</b>	<b>bf: Instance</b>
Publishing organization, volume and year	bf: ProvisionActivity	Identifier	bf: Identifier
File type	bf: FileType	File size	bf: FileSize
Encoding format	bf: EncodingFormat		
<b>Item Level</b>	<b>bf: Item</b>	<b>Item Level</b>	<b>bf: Item</b>
HTML reading, CAJ download, PDF download, etc.	bf: ImmediateAcquisition		

# Other electronic resources의 매핑

- 기타 유형의 전자자원(e-resource)에 대한 BIBFRAME 클래스와 속성 매핑
  - 도서(books), 신문(newspapers), 연감(almanacs), 연구성과(outcomes), 특허(patents), 표준(standards)

**Table 4.** Classes and attributes mapping for other types of electronic resources.

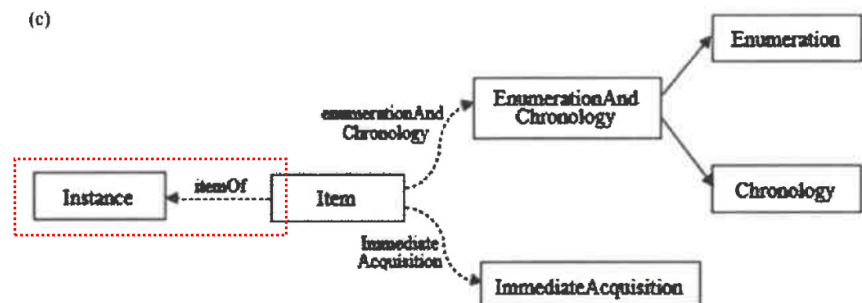
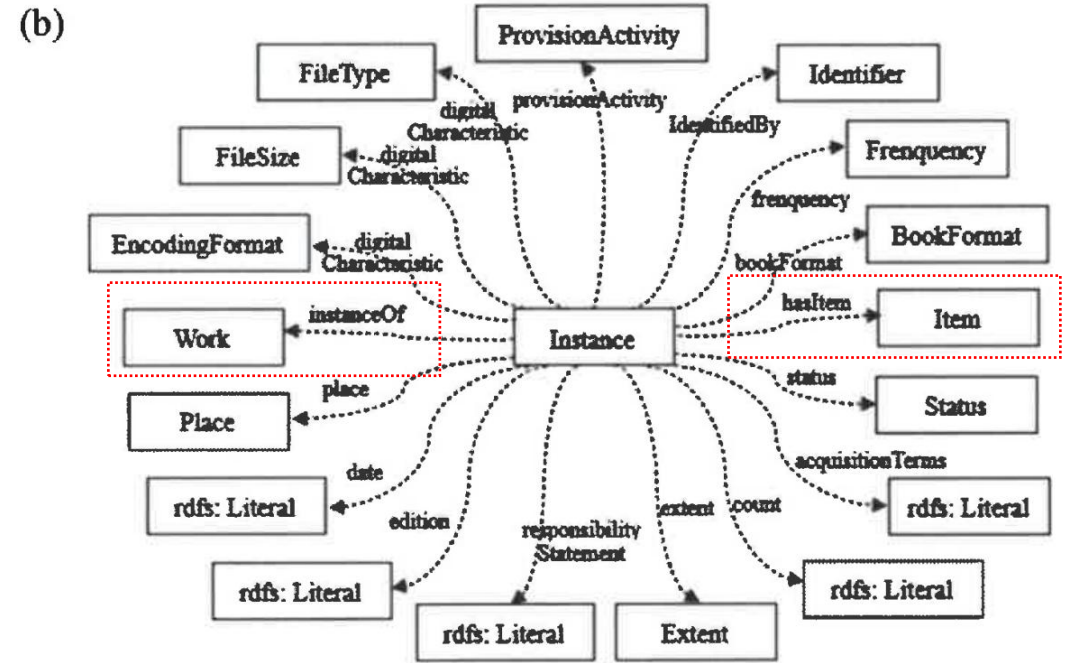
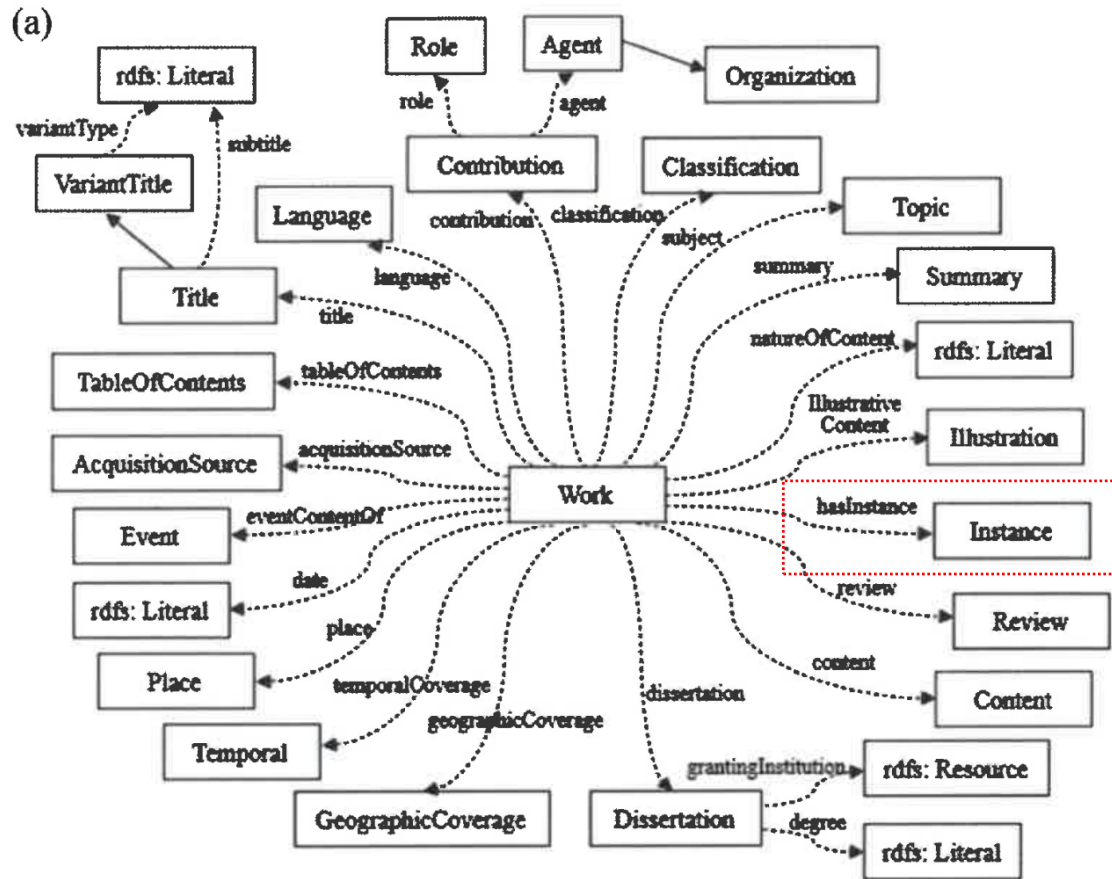
Work Level	bf: Work	Work Level	bf: Work
Title, Title of outcome	bf: Title	Language, Body language	bf: Language
Responsible party, Author, Editor-in-chief, Inventor, Drafter, Accomplisher of outcome	bf: Contribution	Author unit, Applicant, Agent, Drafting unit, Standard technical committee, First completion unit	bf: Organization
Agent	bf: Person	Address	bf: Place
Classification number	bf: Classification	Brief description of content, Brief description of outcome	bf: Summary
Keywords	bf: Topic	Funding	bf: AcquisitionSource
Table of contents of the book, Article catalog, and column catalog	bf: TableOfContents	Introduce an issue	bf: variantType
Subheading	bf: subtitle	Snapshots of the main text, Sovereignty items	bf: illustration
Geography	bf: GeographicCoverage	Patent type, Outcome type	bf: Content
Standardized nature	bf: natureOfContent	Outcome level, Evaluation form	bf: Review
<b>Instance Level</b>	bf: Instance	<b>Instance Level</b>	bf: Instance
Publication information, Distribution units	bf: ProvisionActivity	Identifiers	bf: Identifier
Date of publication, Date of newspaper, Date of entry of results into the database	bf: date	The order in which editions are printed	bf: edition
Date of application, Date of issuance	bf: originDate	Public date, Date of implementation	bf: legalDate
Pages, Total pages	bf: Extent	Number of written characters	bf: count
Statement of responsibilities	bf: responsibilityStatement	Pricing, Paper version pricing	bf: acquisitionTerms
Legal status	bf: Status	File type	bf: FileType
File size	bf: FileSize	Encoding format	bf: EncodingFormat
<b>Item Level</b>	bf: Item	<b>Item Level</b>	bf: Item
HTML reading, CAJ download, PDF download, etc.	bf: ImmediateAcquisition		

## 클래스 및 속성의 확장 및 적용

- 분류(classification) 하위 클래스
  - 주분류(main classification)
  - 중국표준분류(Chinese standard classification)
  - 주제분류(subject classification) 등
- 식별자(identifier)의 하위 클래스
  - ISBN
  - 판차(edition number)
  - DOI
  - 출원번호(application number)
  - 공개번호(disclosure number)
  - 국가 및 지역 코드(country and province code)
  - 표준번호(standard number) 등

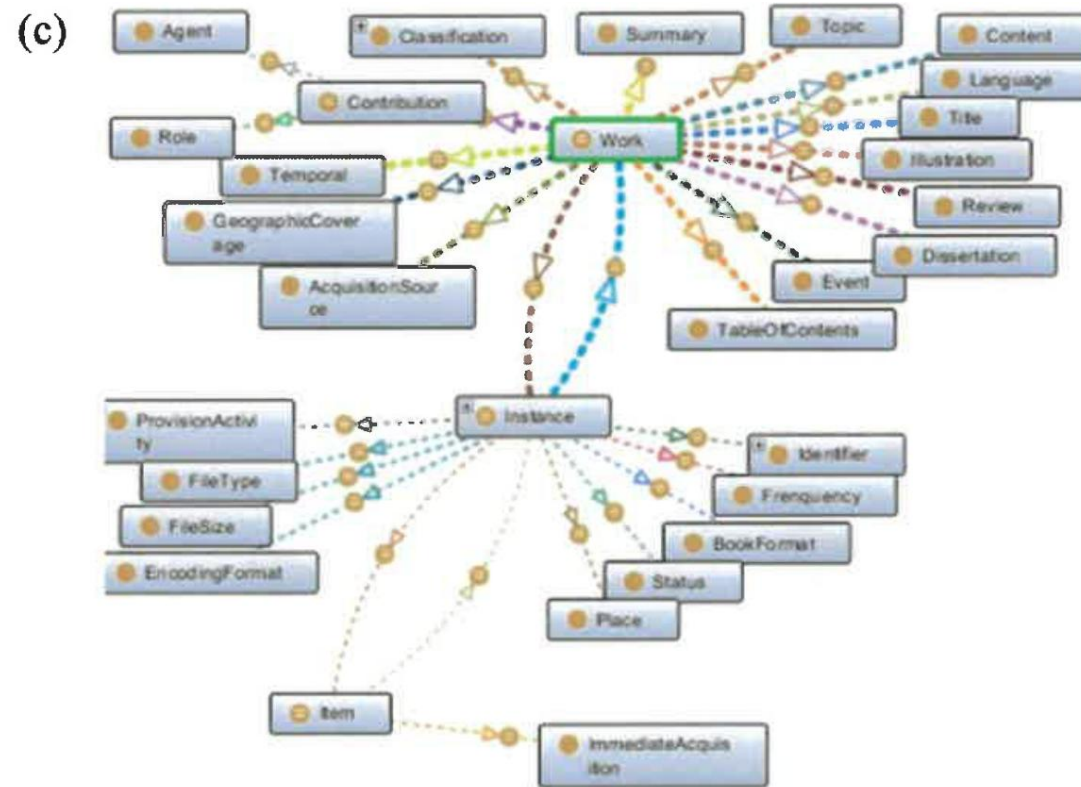
## 다대일 매핑 사례

# 연구내용 ② 전자자원 서지데이터 온톨로지의 관계 구축



# 연구내용 ③ 전자자원 온톨로지 개념모형 구축

- 전자자원 서지데이터 온톨로지의 시각화(visualization), 저장(storage), 질의(query)
- Protégé 소프트웨어를 활용하여 온톨로지 구축



# 시각화(visualization), 저장(storage)

- 출판물(publications) 구축 결과
  - 구축 사례: 「Journal of The Library Science in China」
    - 시각화는 Neo4J 활용

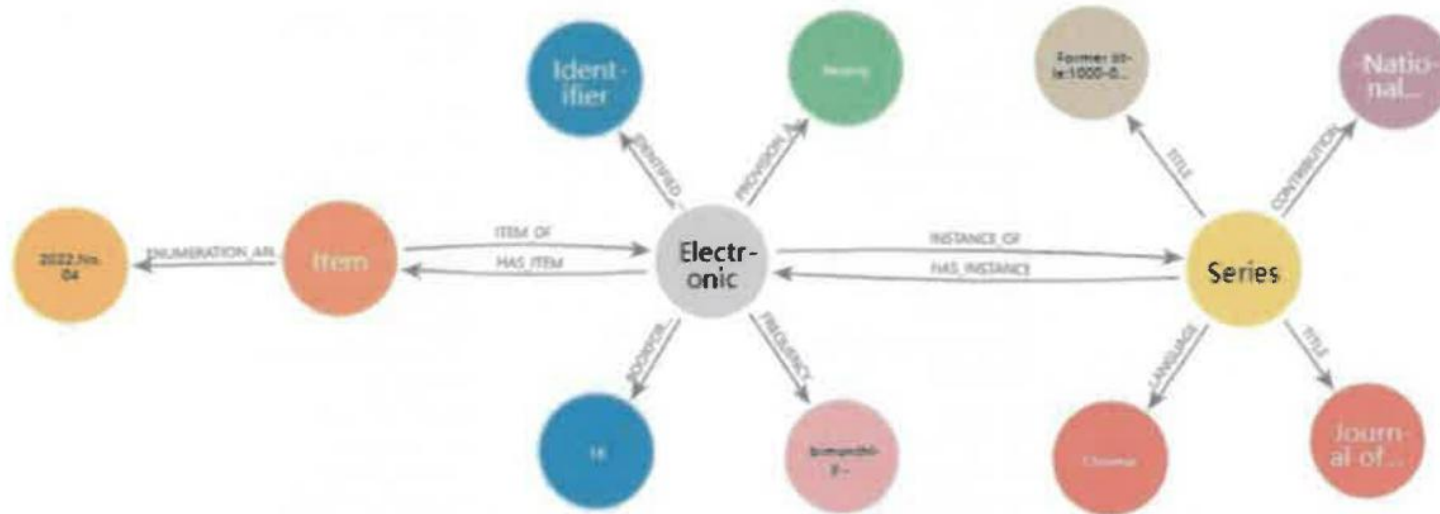


Figure 3. Bibliographic data of Journal of The Library Science in China.

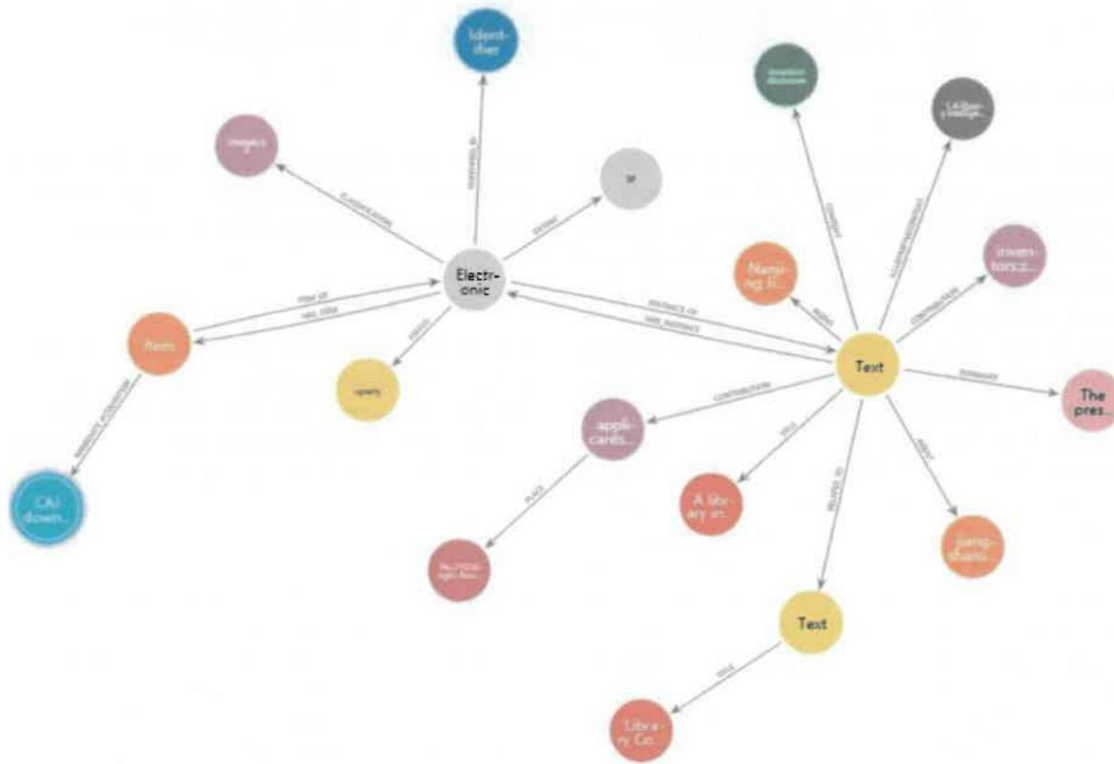
## chronology and enumeration

- “Expected Value: 2022.No.04”
  - Enumeration1: Exploration of the concept of intelligence and its relationship to the information chain and the DIKW chain
  - Enumeration2: New opportunities for the integration and development of public libraries
  - Enumeration3: Analysis of the concept of differentiated thinking based on the perspective of constitutive dimensions



# 시각화(visualization), 저장(storage)

- 기타 유형의 전자자원 중 특허(Patent) 구축 결과
  - 구축사례: 「A library intelligent guidance device and method based on wireless communication sensing technology」



## 관련표준 (연관저록)

- bf:relatedTo

Figure 5. Bibliographic data of a patent.

# 질의(query)

- 질의 사례: 특허의 기여자(contribution)와 대리인(agent), 제목(title), 콘텐츠 유형(content type), 삽화 내용(illustration content)
  - Neo4J의 Cypher 언어를 활용

```

1 MATCH p1=(Work)-[:CONTRIBUTION]→(Contribution1)
2 MATCH p2=(Work)-[:AGENT]→(Agent2)
3 MATCH p3=(Work)-[:TITLE]→(Title)
4 MATCH p4=(Work)-[:CONTENT]→(Content)
5 MATCH p5=(Work)-[:ILLUSTRATIVECONTENT_]→
  (Illustration) RETURN p1,p2,p3,p4,p5
    
```



# 목록 패러다임의 변화

기존	변화
MARC	BIBFRAME
레코드 중심	데이터, 그래프 중심
기술(description)	관계(relation, linking)
목록(catalogue)	지식조직(knowledge organization)
데이터 저장	의미 네트워크

# 발제문

- 저자가 제시한 한계 외에 어떤 연구의 한계가 더 있을까요?
  - RDF 형태로 공개하여 LOD의 가능성을 보여주지 못함
  - 실제 대규모 운영 검증 부재
- 논문에서 제시한 W-I-Item 구조가 전자자원의 특성을 충분히 반영하고 있나요?
  - PDF 다운로드 파일을 Item으로 보는 방식은 적절한가?
  - 스트리밍 콘텐츠나 동적 웹자원도 같은 방식으로 볼 수 있을까? (전자자원의 비물리적 특성)
  - 전통적 장서 개념과 충돌하지는 않는가? (디지털 객체의 표현)
- 논문에서 제시한 관계 속성은 실제 정보검색에서 이용자 탐색 경험을 어떻게 바꿀까요?

# 온톨로지 언어

- 그래프 데이터베이스 (Neo4J)
  - 구조: Node – Relationship – Node
  - 특징: 관계 탐색, 시각화, 연결 구조 표현에 용이
    - 그러나, 의미론적 추론(semantic inference)를 수행하지 못함
- RDF(Resource Description Framework)
  - 구조: Subject – Predicate – Object
  - 특징: URI를 기반으로, 웹 표준에 따라, LD 구현 가능
    - 그러나, 제약 조건을 기반으로 한 추론에 약함
- OWL(Web Ontology Language)
  - RDF에 클래스 정의, 제약조건, 논리적 추론을 더해 의미론적 추론이 가능한 언어