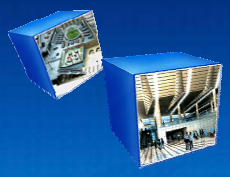
ATOMS 2007

Topic Maps-Driven Semantic Services for National Library of Korea



Presenter Sam Oh

SUNG KYUN KWAN UNIVERSITY

ATOMS 2007

Presentation Outline

- Research Objectives and Scope
- Understanding FRBR Model
- Converting MARC to FRBR
- Converting FRBR to Topic Maps
- Implementing a TM Prototype Enabling FRBR Model

SUNG KYUN KWAN UNIVERSITY

Research Objectives and Scope

- Establishing resource description structure based on international standard
- Making resource management smart and systematic
- Towards semantic DL service via user participation

- *The FRBR model is revolutionary. The (computer) catalogue is not seen as a sequence of bibliographic records and a replica of the traditional card catalogue, but rather as a network of connected data, enabling the user to perform seamlessly all the necessary functions.*

- Dr. Maja Žumer. National and University Library, Ljubljana, Slovenia

FRBR Model...

- FRBR = Functional Requirements for Bibliographic Records
- Developed by cataloging experts working under the auspices of IFLA (International Federation of Library Associations and Institutions)
- FRBR is a conceptual model (not a standard!)
  - FRBR systematically models the bibliographic universe

Impetus: a changing environment

- Automated systems for creation & processing of bibliographic data
- Growth of large-scale databases (national & international)
- Shared cataloging programs
- Economic pressure to reduce cataloging costs
- Adapting practice to accommodate new formats & a networked environment
- Changing user expectations and needs

Value of FRBR

- FRBR study process and report:
  - Fresh look at what functions bibliographic records perform
  - FRBR systematically models the bibliographic universe
- Helps us answer important questions:
  - What information is of the most value to users of the catalog?
  - How can that most valuable information be used more effectively?
- Promises to inspire next generation catalogs and DL systems.
  - Illuminates user tasks
  - Clarifies how catalogs and DLs should function

FRBR Study

Primary objectives of the FRBR study:

- Provide a clearly defined, structured framework for relating bibliographic data to user needs
- Recommend a basic level of functionality for records created by national bibliographic agencies

Modeling the bibliographic universe:

- Utilizes an entity-relationship framework:
  - Entities (a class of things)
  - Relationships (associations among entities)
  - Attributes (characteristics of the entities)

What FRBR says USERS do:

Four User Tasks:

- **Find** entities that match the search criteria
- **Identify** entities (confirm that the user has found what they were looking for)
- **Select** an entity that matches the user's criteria
- **Obtain** access to the entity through purchase, loan, or through electronic remote access

And, maybe a fifth:

- **Navigate**

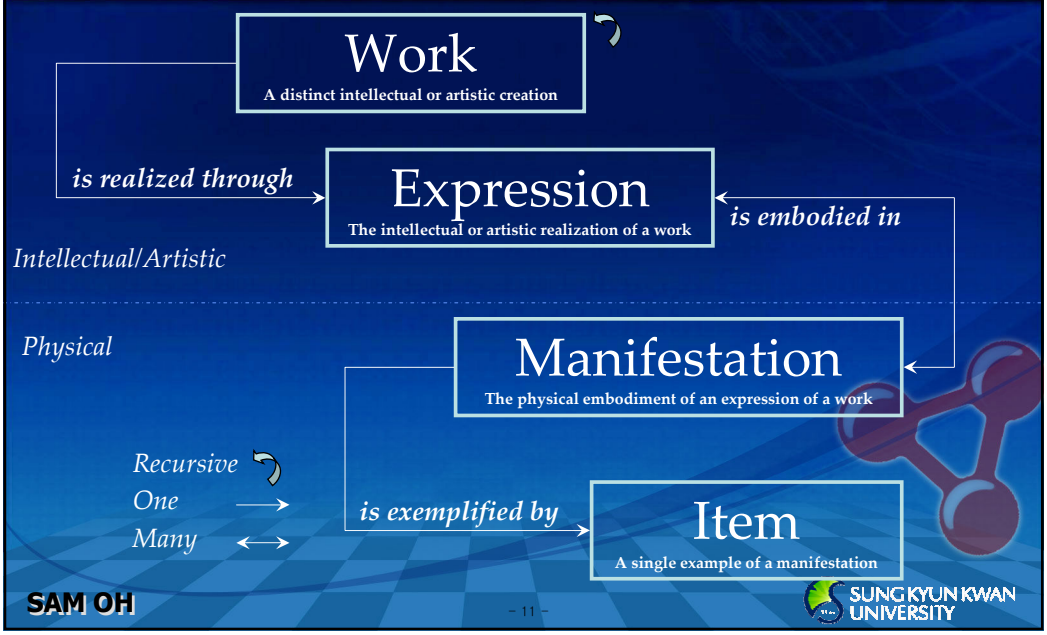
FRBR 3 Groups of Entities

- **Group 1**
  - **products of intellectual or artistic endeavor that are named of described in bibliographic records**
    - work, expression, manifestation, item
- **Group 2**
  - **entities responsible for the intellectual or artistic content, the physical production and dissemination, or the custodianship of such products**
    - person, corporate body
- **Group 3**
  - **entities that serve as the subjects of intellectual or artistic endeavor**
    - concept, object, event, place

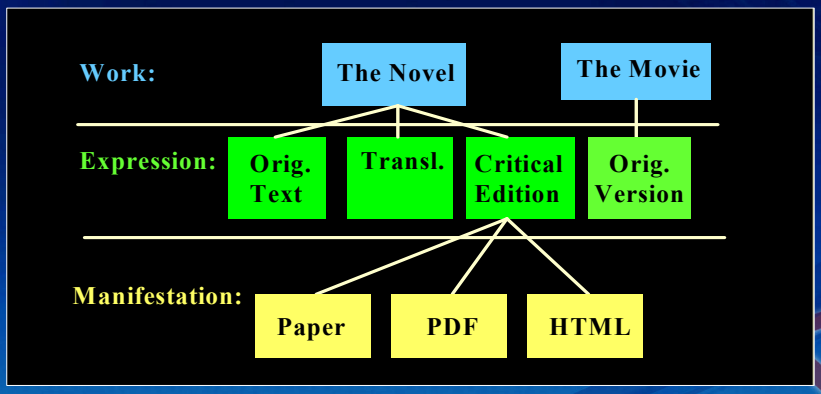
FRBR Group 1

- Work
  - A distinct intellectual or artistic creation
- Expression
  - The intellectual or artistic realization of a work
- Manifestation
  - The physical embodiment of an expression of a work
- Item
  - A single example of a manifestation

FRBR Group 1…



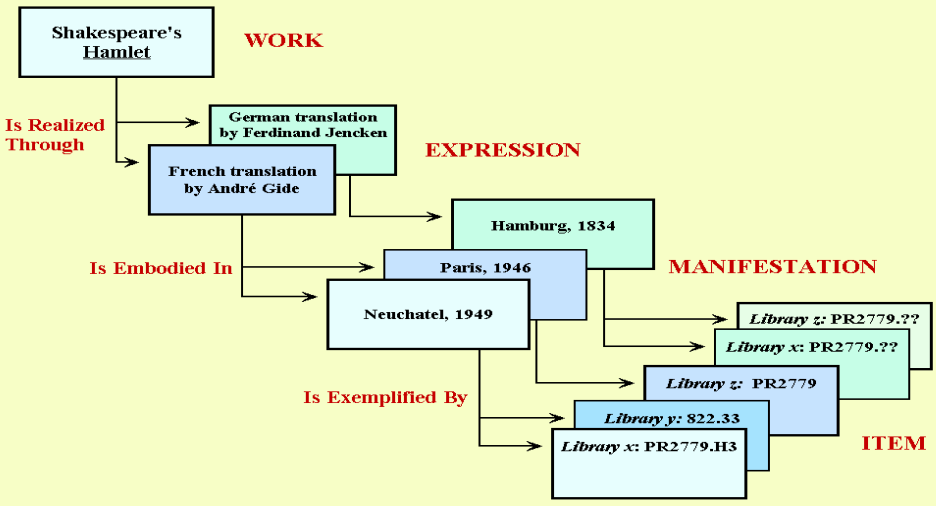
Group 1



From: Tillett, Barbara "AACR2's Strategic Plan and IFLA Work towards an International Cataloguing Code" (2002)

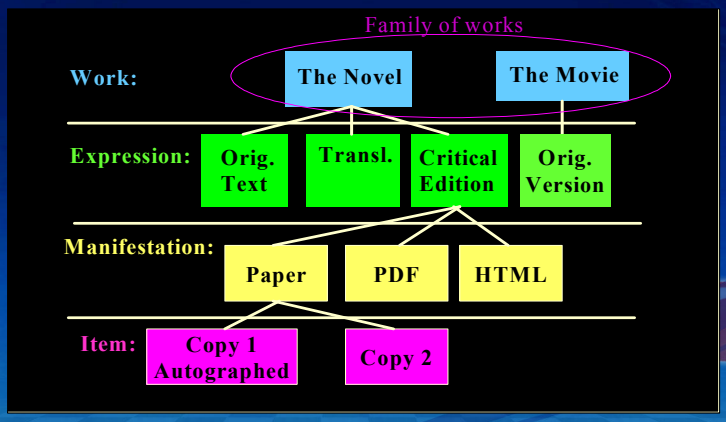
ATOMS 2007 Topic Maps-Driven Semantic Services For National Library of Korea
FRBR Entity Levels

4-level FRBR Model



FRBR Entity Levels

Group 1



From: Tillett, Barbara "AACR2's Strategic Plan and IFLA Work towards an International Cataloguing Code" (2002)

Group 1 Attributes

- Work
  - Work title, form or genre, date, performance medium, intended audience
- Expression
  - Expression title, form of the expression, language of the expression, type of score, scale of a map
- Manifestation
  - Manifestation title, publisher, date of publication, form of carrier, dimensions, manifestation identifier (e.g. ISBN), terms of availability
- Item
  - Location or call number, barcode, provenance, condition, access restrictions on an item

Group 2 – Responsible Entities

- Entities responsible for the intellectual or artistic content, the physical production and dissemination, or the custodianship of such products

Group 2 entities:

- Person
- Corporate Body

And perhaps a third …

- Family

Group 2 Attributes

- **Person**
  - Names, dates, titles or other designations
- **Corporate body**
  - Name, number, place, date, other designation

Group 3 - Subject Entities

- Entities that serve as the subjects of intellectual or artistic endeavor

Group 3 entities:

- **Concept:** topical subject heading
- **Object:** name for an object
- **Event:** name for an event
- **Place:** name for a place
- Plus the Group 1 and 2 Entities:
  - Works about Works/Expressions/Manifestations/Items
  - Works about Persons and/or Corporate Entities

Group 3 Attributes

- Attributes of Group 3 entities:
  - Concept
    - term
  - Object
    - term
  - Event
    - term
  - Place
    - term

Relationships

- Link entities to one another
  - Entities of different groups (person and work)
  - Entities of the same group (work and expression)
- Collocation and navigation

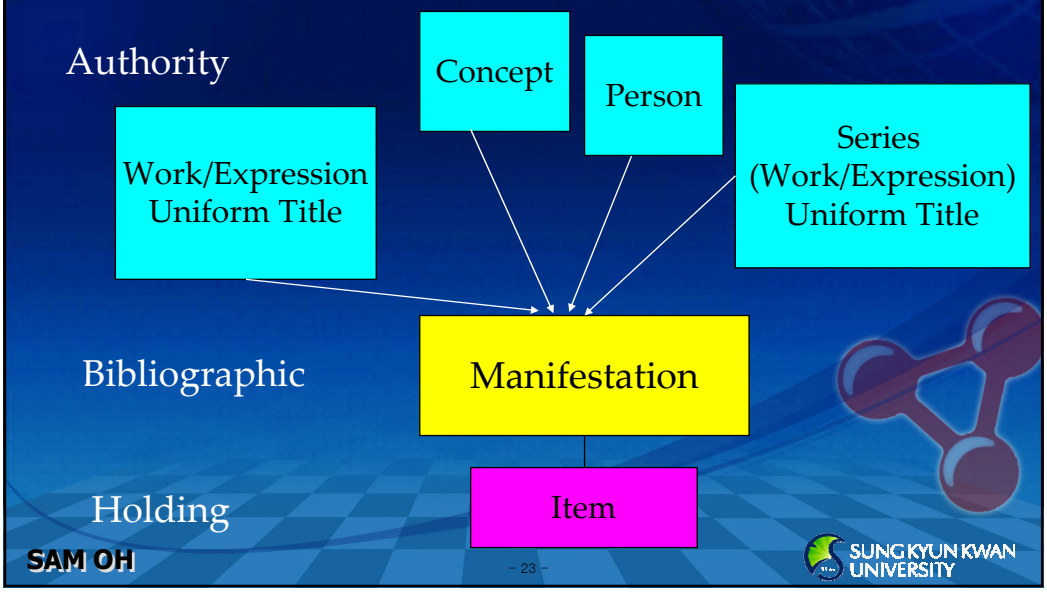
Relationships Between FRBR Groups

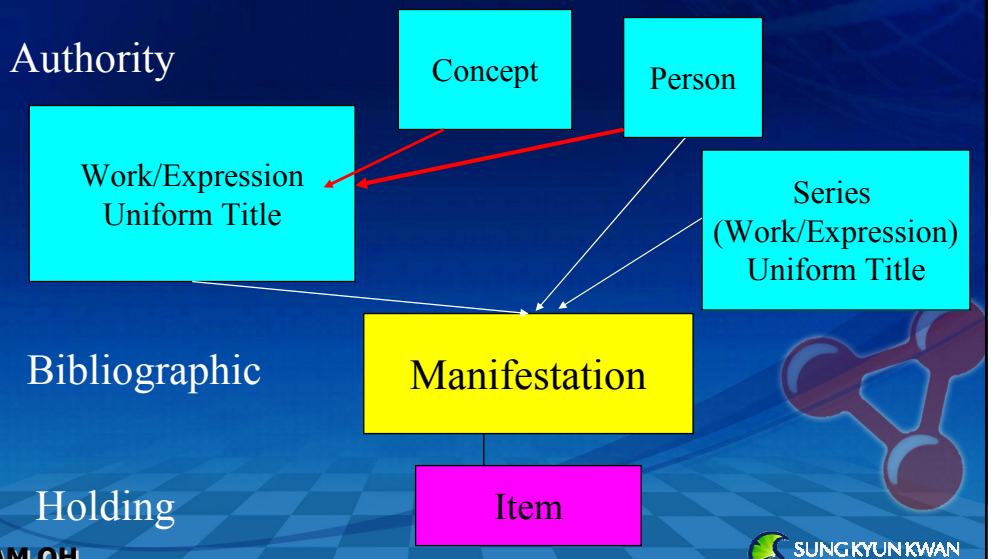
- Work ⇔ Person
  - Created by
- Expression ⇔ Person
  - Translated by
- Manifestation ⇔ Corporate body
  - Produced by
- Item ⇔ Corporate body
  - Owned by
- Concept ⇔ Work
  - Is subject of

Relationships within a group

- Work to work
  - Successor
  - Supplement
  - Complement
  - Summarization
  - Adaptation
  - Transformation
  - Imitation
  - Whole/part

Scenario A (Now)





➢ Incorporating the concepts of the FRBR model in systems:

- **Superior presentation of search results**
  - Esp. in large files – more intuitive clustering
- **May help streamline library cataloging**
  - Reduces repeated keying of work-related info
- **Bibliographic & management intelligence**
  - New insights into works (e.g., OCLC's 1000 list)
  - Libraries can operate at workset level (e.g., ILL)

- OCLC Research:
  - Top 1000 titles http://www.oclc.org/research/top1000
  - xISBN http://www.oclc.org/research/researchworks/xisbn/
  - Fiction Finder http://fictionfinder.oclc.org/
  - Curiouser http://www.oclc.org/research/projects/curiouser
- OCLC production services
  - Open WorldCat http://www.oclc.org/worldcat/open
  - FRBR-inspired FirstSearch WorldCat (coming soon)

Mapping Between FRBR Attributes & MARC21 Elements

Entities	Attributes	MARC21 Elements	MARC21 Description
Work	created by	100/110/111 700/710/711 $t,$n,$p,$d,$f,$k,$s,$m,$r,$l,$o,$h	person/corporate body
	Title of work	130 $a,$n,$p 240/243/245 $a,$n,$p	Uniform title Collective uniform title Title statement
	Date of work	130 $d 240/243 $d 245 $g	Date of treaty signing
	Form of work	130 $k 240/243/245 $k	Form subheading
	Medium of performance	130 $m 240/243 $m	Medium of performance for music
	Key	130 $r 240/243 $r	Key of music
	Coordinates	255 $c	Statement of coordinates

Mapping Between FRBR Attributes & MARC21 Elements

Entities	Attributes	MARC21 Elements	MARC21 Description
Expression	Form of expression	leader/06	Type of record
	Language of expression	130 $l 240/243 $l 008/35-37	Language of a work language
	Other distinguishing characteristic	130 $o 240/243 $o 245 $s	Arranged statement for music Revision
	Type of score	254 $a Music 008/20	Musical presentation statement Format of music
	Scale	255 $a 034 $b,$c,$h	Statement of scale

ATOMS 2007
Topic Maps-Driven Semantic Services For National Library of Korea
Converting MARC to FRBR

Mapping Between FRBR Attributes & MARC21 Elements

Entity	Attribute	MARC21 Elements	MARC21 Description
Manifestation	Title of manifestation	245 $a,$n,$p 242 $a,$n,$p	Title Number of part Name of part
	Statement of responsibility	245 $c 242 $c	Statement of responsibility
	Edition/issue designation	250 $a,$b	Edition statement Remainder of edition statement
	Numbering	362 $a	Dates of publication and sequential designation
	Publisher/distributor	260 $b 533 $c	Name of publisher, distributor Agency responsible for reproduction
	Date of publication/distribution	260 $c 533 $d	Date of publication, distribution Date of reproduction

ATOMS 2007

Topic Maps-Driven Semantic Services For National Library of Korea

Converting MARC to FRBR

Mapping Between FRBR Attributes & MARC21 Elements

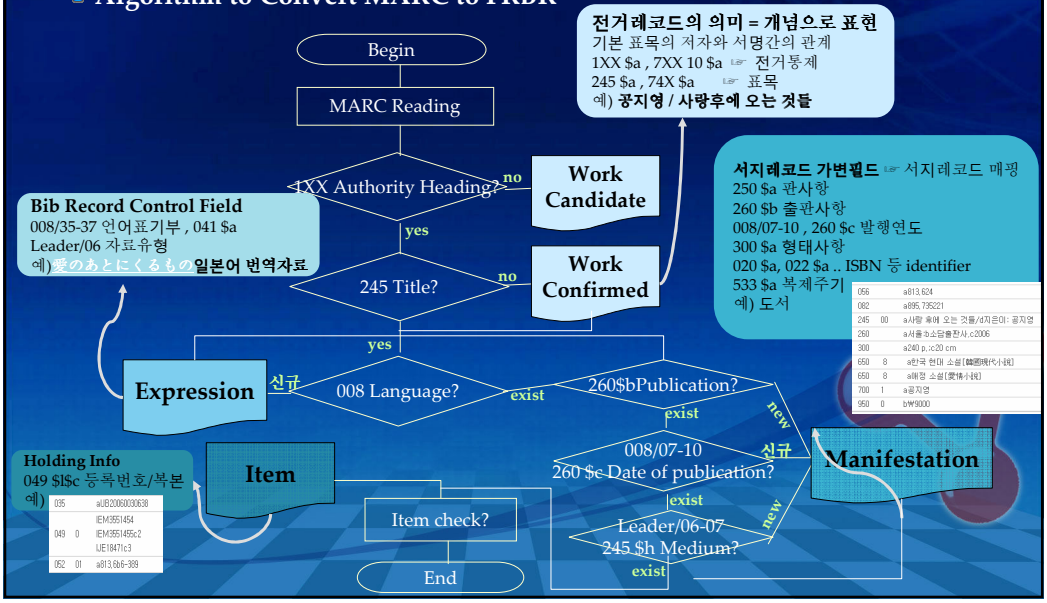
Entity	Attribute	MARC21 Elements	MARC21 Description
Manifestation	Form of carrier, extent of carrier	300 $a 533 $e	Extent Physical description of reproduction
	Dimensions of carrier	300 $c 533 $e	Dimensions Physical description of reproduction
	Physical medium, capture mode, playing speed, groove width, kind of cutting, tape configuration, kind of sound, special reproduction characteristics, color, polarity, presentation format, file characteristics	300 $b	Other physical details
	Reduction ratio	340 $f 533 $e(?)	Production rate/ratio Physical description of reproduction
	Manifestation identifier	BK:020,027,074 SE:022,030,074 Other:024,028,074	ISBN,STRN:BK, ISSN, Coden:SE Other standard identifier, publisher number:other, Form of musical composition code:074t

ATOMS 2007

Topic Maps-Driven Semantic Services For National Library of Korea

Algorithm to Convert MARC21 to FRBR

Algorithm to Convert MARC to FRBR



An Algorithm to Convert MARC21 to FRBR

1. Read MARC record one by one.
2. Read 1XXtag about author and apply normalization rules. (100/110/111 tag $abcdq extract text)
3. If the author already exists in the DB, extract the title. If not, that becomes a new candidate for a new work.
4. In the 245tag $a, extract the title and compare it to see whether it matches one of works in the DB.
   1) Apply normalization rules to extracted title.
   2) If the title in 245tag does not match, extract the text from 740tag $a or 507tag $t and compare it with the existing works.
   3) When you find a title that matches a work, extract the text in 245tag $a as an expression of the work and move to the expression routine.
   4) Fix the author and title as work, then move to expression module.

An Algorithm to Convert MARC21 to FRBR…

6. Check the language, 008tag / 35-37, if that is a new one, register it as a new expression.
   - If 041tag exists, separate the text from $a and $bch, the $bch becomes an expression.
7. Describe expression concept as author, title, language, and type.
8. After you went through work and expression steps, check manifestation items.
   - After you extract a publisher info from 260tag $b, apply normalization rules, then check whether it matches existing ones.
   - Extract publication date from 008tag / 07-10 and check it matches with others.
   - Treat the value of leader/06-07 as upper classification category, if 245tag $h exists, separate resource type from it and check whether it matches with others.
9. Registration of a manifestation will be done per record and manifestation. If that is an item of manifestation, it should be excluded.
10. Extract holding information from items and register it.

Connecting FRBR with Cataloging Items

저작
Author: Scott, Walter, Sir, 1771
Work: Heart of Midlothian

표현형 1
Form: text - English

구현형 1
[Edition]
- Title: The heart of Midlothian
- Statement of responsibility: Sir Walter Scott ; edited with an introduction and notes by Claire Lamont.
- Imprint: Oxford University Press, 1999
- Physical Description: xxviii, 583 p. : ill. ; 20 cm.
- ISBN: 019283567X

구현형 2
[Edition]
- Title: The heart of Mid-Lothian
- Statement of responsibility: Sir Walter Scott ; edited with an introduction and notes by Tony Inglis.
- Imprint: Penguin Books, 1994
- Physical Description: lvi, 793 p. ; 20 cm.
- ISBN: 0140431292

표현형 2
Form: text - Russian

구현형 1
[Edition]
- Title: Edinburgskaia temnitsa
- Imprint: 1957
- Physical Description: 630 p. illus. 21 cm.

Input (marc reading)

레코드1

리더	00572nam 2200205 k 4500	
TAG	IND	내용
001		KMO199509675
005		19950615135518
008		950615s1994 ulk 000 f kor
040		a011001c011001
049	0	lEM12576O5v1994
052	01	a813.6b공784ㅁc1994
056		a813.623
082		a895.734219
100	1	a공지영
245	10	a무소의 불처럼 혼자서 가라:c공지영장편소설/d공지영 지음
260		a서울:b문예마당,c1994
300		a304p.;c23cm
440	00	a문예마당 소설선:v3

레코드2

리더	00935nam 2200277 k 4500	
TAG	IND	내용
001		KMO199941795
005		20000103193550
008		991126s1999 ulk 000 j kor
020		a899436646g03810
049	0	lEM1995821 lEM1995823c2
052	01	a813.6b공784ㅁc
056		a813.624
082		a895.734221
100	1	a공지영
245	10	a존재는 눈물을 흘린다:c공지영 소설/d공지영 지음
260		a서울:b창작과비평사,c1999
300		a310p.;c23cm
505	00	t랑기의 역사.- t고독.- t길 - t조용한 나날.- t진지한 남자.- t모스끄바에는 아무도 없다
653		a존재ㅣb눈물ㅣa공지영ㅣa소설ㅣa랑기ㅣa역사ㅣa고독ㅣa길ㅣa진지ㅣa남자ㅣa모스끄바
740	0	a랑기의 역사
740	0	a고독
740	0	a길
740	0	a조용한 나날
740	0	a진지한 남자
740	0	a모스끄바에는 아무도 없다
950	0	b₩7500

레코드3

리더	00954nim 2200253 k 4500	
TAG	IND	내용
001		KMU200200247
005		20020521112631
007		sdufsngnumnued
008		020514e2001 ulnnnn f kor
020		a8989556147(1(1))g08810:b₩12000
020		a8989556147(1(2)g08810
020		a8989556147(2)g08810
049	0	lAV364103v1(1) lAV364104v1(1):c2 lAV364105v1(2) lAV364106v1(2):c2 lAV364107v2 lAV364108v2c2ISD
052	01	a813b4c1(1)-2
056		a81324
245	00	a존재는 눈물을 흘린다 h[녹음자료] : c공지영 소설, n1(1) - 2/ d공지
260		a서울:b소리공화국,c2001
300		a음반 2매, 녹음 카세트 1개:bCD지혈, 스테레오:c12 cm+e녹음 카세트 1
653		a공지영 a존재a눈물a소설
700	1	a공지영
700	1	a성연미
700	1	a홍승성
700	1	a석원희
950	0	b₩12000c(set 가격)

레코드4

리더	00614ndm 2200205 k 4500	
TAG	IND	내용
001		KMU000099056
005		20041129125752
007		ta
008		041112q19601997ko kvd n kor
049	1	lAV0000639787NB
052	01	a673.51b97280
056		a673.5124
245	00	a무소의 불처럼 혼자서 가라 h[악보]/ d정영아 f曲: e공지영 f詩
260		a[발행지불명]:b[발행자불명],c[1960-1997]
300		a악보 1매;c26 cm
500	00	a심의번호 : 3644
653		a대중음악a가요a무소a불a혼자
700	1	a정영아
700	1	a공지영
950	1	a가격불명

레코드5

리더	00983ngm	
TAG	IND	내용
008		050701s2
049	0	lAV69883 lAV69883 lAV69883 lAV69883
052	01	a687.21b39c16(1)-16(2)
056		a687.2124
245	00	aTV문학관 h[비디오 녹화자료], n16(1) - 16(2), p누구에게나 마음속의 강물은 흐른다/dKBS미디어 [편]
246	1	i음기등표제:aKBS VCD TV문학관
260		a서울:bKBS미디어,c2005
300		a비디오디스크 2매(90분):b천연색, 유성:c12 cm
500	10	d원작자: 공지영
538		aVideo-CD
653		aTV문학관a강불a소설a문학a드라마
700	1	a공지영
710		aKBS미디어
856	42	uhttp://www.kbsmedia.co.kr
910	0	a케이비에스미디어
940		aE1브이 문학관
940		aKBS TV 문학관
950	0	b₩19250

Advantages of Topic Maps in Providing Semantic Services

- Provides excellent collocation capability – Katz's second purpose of cataloging
- TM implementation is relatively easy compared with RDF/OWL so one can expect a better return on investment.
- Excellent candidate in modeling FRBR-based data and provides great interoperability with other semantic data.
- You can easily change your ontology schema without modifying applications, which offers excellent advantage in terms of management cost.
- Provides flexible ways to model ontology.
- Provides flexible ways to convert any structure data, which makes it easy to utilize legacy data.
- Integrating TM with Web2.0 technology is easy and natural.

Refining MARC-Ontology Using Topic Maps

1. **Topic Types**
   - **Among FRBR entities from group 1,2,3 and MARC metadata elements, select the elements that can create "useful" collocation, assign them as topic types.**
   - **Map FRBR model entities into MARC metadata elements and establish them as topic types.**
2. **Association Types**
   - **Establish association types between topic types.**
   - **Apply FRBR relationships as TM associations between topics and also establish associations between metadata elements.**
3. **Occurrence Types**
   - **Except for MARC metadata elements that were converted as topic types, the remaining MARC elements will be defined as TM occurrence types.**
   - **Classify occurrence types based on element values of FRBR model and assign them to appropriate topic types.**
   - **Since MARC records are usually based on manifestation, most MARC metadata elements and other occurrence types are assigned to manifestation.**

Defining Topic Types

Topic Types

Expand all nodes

- 개별자료
- 구현형
- 국가
- 단체
- 대상
- 매체유형
- 발행일
- 발행처
- 배경
- 분류체계
- 사건
- 언어
- 이용대상
- 인물
- 자료유형
- 장르
- 장소
- 저작
- 주제
- 표현형

Topic Types	FRBR Group & MARC Element
Work/Expression Manifestation/Item	FRBR Group 1, MARC Title
Person / Corporate Body	FRBR Group 2, MARC Personal name, Corporate Name
Object / Event / Place	FRBR Group 3
Country / Genre	MARC Subject (KORMARC 650tag)
Type	MARC leader's type of record (Expression Criterion)
Format	MARC Medium (245 $h, Manifestation Criterion)
Language	MARC Language statement (Expression Criterion)
Classification	MARC KDC classification number
Context	FRBR Model의 개체 속성정보
Publisher, Date	MARC publication
Audience	MARC Target audience / User tagging info
Subject	User subject tagging info

ATOMS 2007
Topic Maps-Driven Semantic Services
For National Library of Korea

KORMARC Elements and TM Occurrence Types

KORMARC Elements	Topic Maps Occurrence Types
Title and related fields (1XX/24X/4XX/7XX)	Other words associated with a name, Remainder of title, Parallel title, Original title, series statement, uniform title
Publication (26X)	Edition statement, Place of publication, Frequency, Regularity
Physical description (3XX)	Extent, Other physical details, Dimensions, Accompanying materials, size of unit
Note (5XX), Subject(653)	Contents note, Restrictions on access, Immediate source of acquisition, keyword
Non-book bibliographic Note(5XX)	System details note, Music format, Participant or performer note, Exhibitions note, Date/time and place of an event note, Creation/production credits note, scale note
Control field (0XX)	DDC, KDC, ISBN, Price, Korean National Library call number, identifier number, language note
Other service Note	URL, book review, status of book circulation, abstract, image of title

Occurrence Types by Topic Types

Topic Type	Occurrence Types
Work	uniform title
Expression	Original title, language note, Frequency, Regularity, Restrictions on access
Manifestation	Contents note, Immediate source of acquisition, Extent, Other physical details, Dimensions, Accompanying materials, size of unit, Edition statement, System details note, Music format, Participant or performer note, Exhibitions note, Date/time and place of an event note, Creation/production credits note, scale note, Other words associated with a name, Remainder of title, Parallel title, DDC, KDC, ISBN, Price,
Item	Korean National Library call number, identifier number
Publisher	Place of publication,
Classification	KDC, DDC
인물	Dates associated with a name?
Subject	keyword

ATOMS 2007 Topic Maps-Driven Semantic Services For National Library of Korea

Defining Association Types

Topic Type	Association Type	Topic Type
Work	Work form	Genre
	Realized Through	Expression
	Created by	Person/ Corporate Body
	Context for	Context
	Work classified	Classification

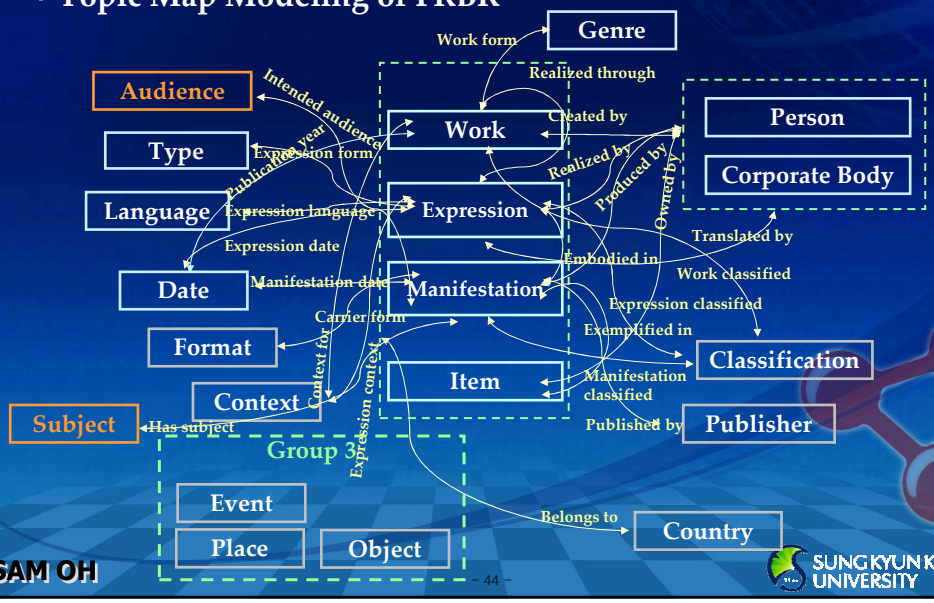
Defining Association Types

Topic Type	Association Type	Topic Type
Expression	Embodied in	Manifestation
	Realized by	Person/ Corporate Body
	Expression form	Type
	Expression language	Language
	Expression classified	Classification
	Expression context	Context
	Expression date	Publication Date
	Translated by	Person/ Corporate Body

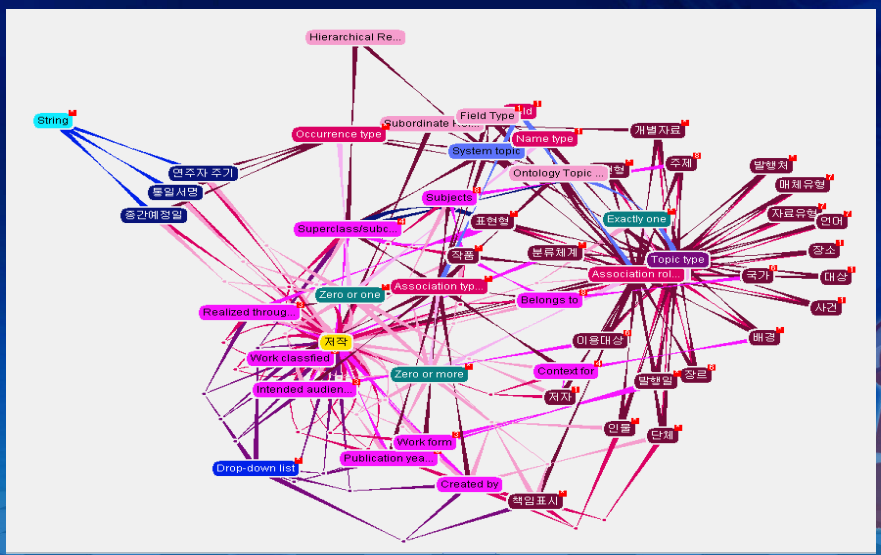
Defining Association Types

Topic Type	Association Type	Topic Type
Manifestation	Exemplified in	Item
	Produced by	Person/ Corporate Body
	Carrier form	Format
	Published by	Publisher
	Manifestation classified	Classification
	Intended audience	Audience
	Manifestation date	Publication Date
Subject	Has subject	Work Expression Manifestation
Country	Belongs to	

Topic Map Modeling of FRBR



Graphical TM View of MARC



ATOMS 2007
Topic Maps-Driven Semantic Services
For National Library of Korea
OCLC Fiction Finder
Gone with the wind.
Mitchell, Margaret, 1900-1949
Work
293 editions
27 languages
11291 libraries
Item
Summary:
Genres: Historical fiction | War stories | Love stories | Epic fiction
Characters: O'Hara, Scarlett (Fictitious character)
Settings: 1861 - 1865 | United States | Georgia | 1800 - 1899
Subjects: American Civil War, 1861-1865 | Women | Plantation life
Audience:
Kids
General
Special
Expression
Manifestation
Editions
Genres
Characters
Settings
Subjects
Title / Author
OCLC #
Date
Format
Aud.
1. Gone with the wind / by Margaret Mitchell 00290058 1936 Book General
2. Gone with the wind 07909619 1936 Book General
3. Gone with the wind / by Margaret Mitchell 35240399 1964 Book General
4. Gone with the wind 01279858 1973 Book General
5. Gone with the wind / Margaret Mitchell 28491920 1936 Book General
6. Gone with the wind / Margaret Mitchell 25412027 1992 English Large print General
7. Gone with the wind / Margaret Mitchell ; with a new preface by Pat Conroy and an introduction by James A. Michener 33947306 1996 English Book General
8. Gone with the wind / Margaret Mitchell 02931402 1936 English Book General
9. Gone with the wind. Illustrated by Ben Stahl 00878256 1961 English Book General
10. Gone with the wind Margaret Mitchell 45950921 2001 English Book General

Main Screen

ATOMS 2007

Topic Maps-Driven Semantic Services
For National Library of Korea

NLK Prototype System

Display of Literary Works

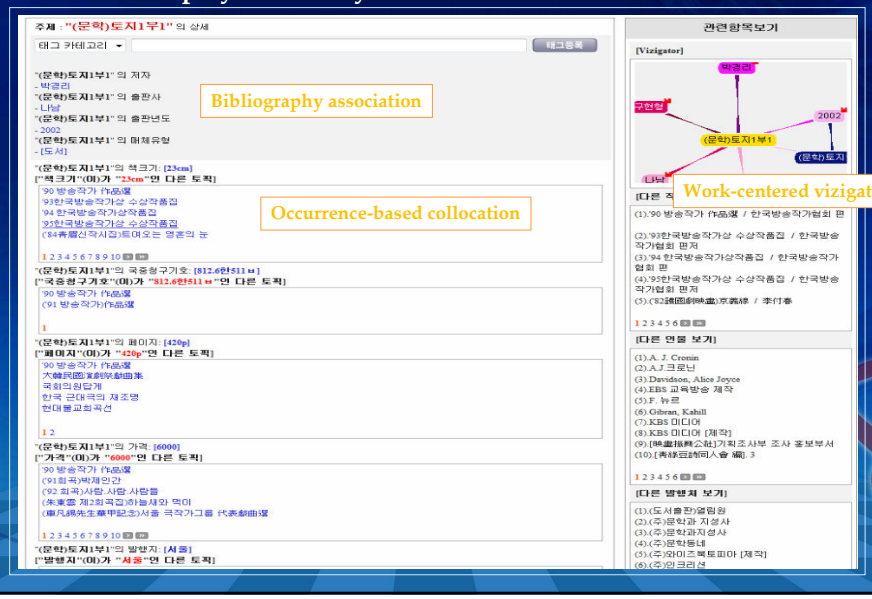
ATOMS 2007

Topic Maps-Driven Semantic Services
For National Library of Korea

NLK Prototype System

Display of Authors

Detailed Display of Literary Works



➢ Future Research Project Agenda

- **National Library of Korea**
  - Designing Semantic Digital Library Architecture
- **National Archives of Korea**
  - Designing Semantic Archiving Retrieval System
- **National Assembly Library**
  - Designing Smart and Interoperable Metadata System

Q & A

Discusstion