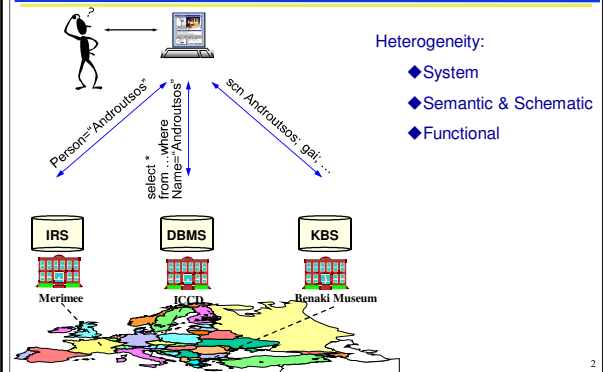


Declarative Specification of Z39.50 Wrappers using Description Logic

Y. Velegrakis

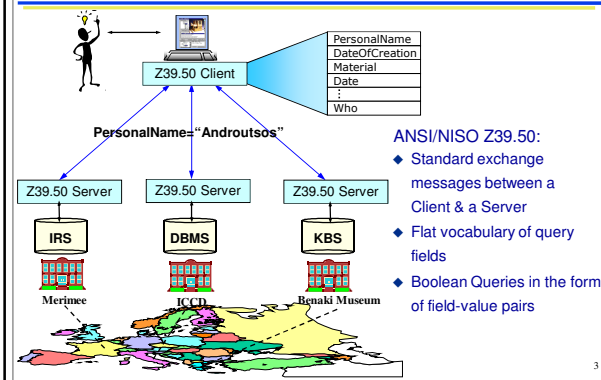
1

Accessing Distributed/Heterogeneous Info.



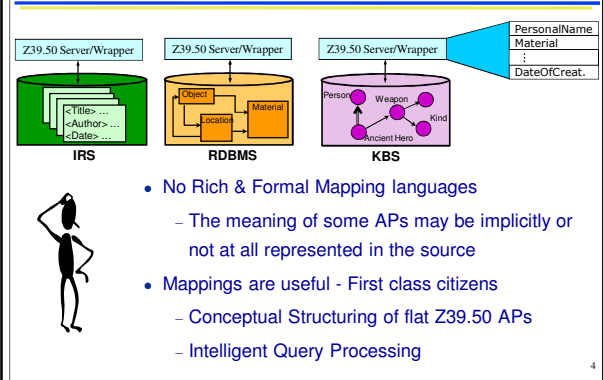
2

Z39.50 for Digital Libraries and Museums



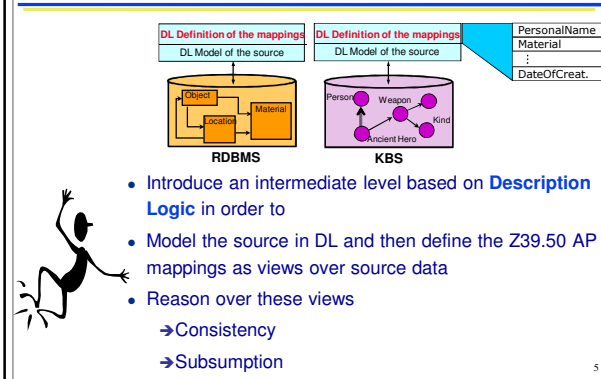
3

Z39.50 Wrapping Issues



4

Our Approach



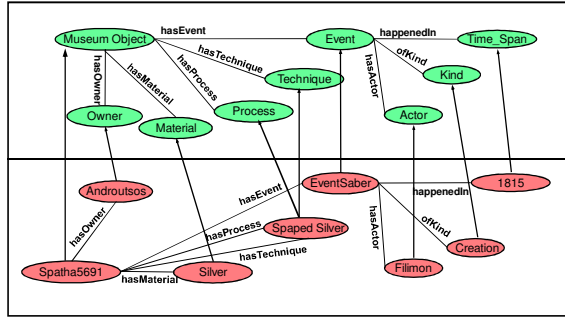
5

Outline

- Represent the source in DL
- Define the AP mappings within the DL framework
- Describe the advantages of that way of mapping
- Explain the query processing in DL
- Describe the benefits of the use of DL for the Z39.50 wrapper specification
- Conclusion

6

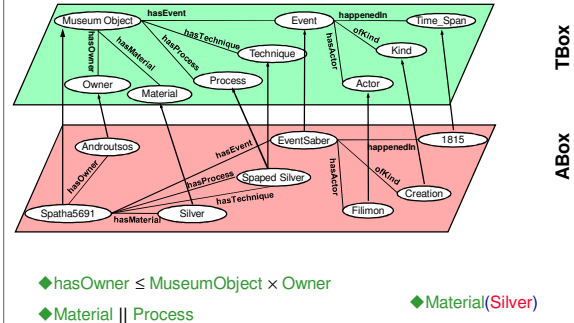
An Example of a Cultural Information Source



Aquarelle and CimiZit Projects

7

Viewing Source Data and Schema in DL



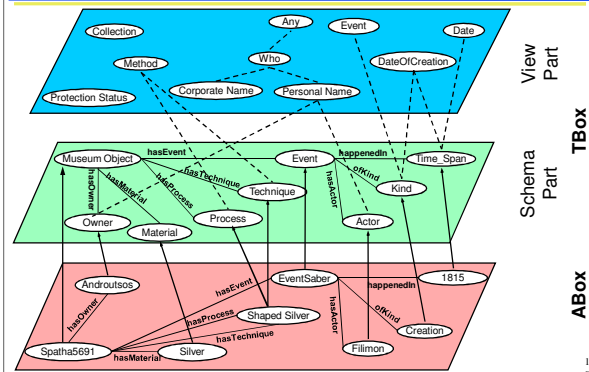
8

Defining the Access Point Mappings in DL

- **Date** = **Time_Span**
- **PersonalName** = **Actor** \cup **Owner**
- **Who** = **PersonalName** \cup **CorporateName**
- **Any** = **Who** \cup **What** \cup **When** \cup **Where**
- **Method** = **Material** \cap **Process**
- **DateOfCreation** = $\exists(\text{happenedIn})^{-1}.(\exists \text{offKind}.\{\text{"Creation"}\})$

9

Putting All Together



10

Ensuring the Z39.50 Wrapping Quality

- Represent even unsupported APs, i.e., not mapped in the source
ProtectionStatus = \perp
- Map APs whose meaning is only implicitly represented
Collection = $\{\text{"Benaki Gun Collection"}\}$
- Allows the formal validation of ill defined mappings by the use of
DL Consistency Checking.
Method = **Material** \cap **Process** $\rightarrow \Sigma \models \text{Method} \equiv \perp$

11

Z39.50 Query Translation to DL

- Q: (**PersonalName**=**"A"**, **Truncation**=**"Right"**) **AND** **Date**=1821
- For simple queries (no qualifiers) the **DL Instance Checking** reasoning service is used, e.g. $Q_1: \text{Date}=1821 \rightarrow \text{Date}(1821)$
 - Truncation/Relation qualifiers are expressed with **Test Functions**
 $Q_2: (\text{PersonalName}=\text{"A"}, \text{Truncation}=\text{"Right"}) \rightarrow (\text{PersonalName} \cap \text{TEST-C}(\text{rtrunc}_A))$
Note: Unsupported test functions are mapped to false
 - **AND**, **OR**, and **AND-NOT** are translated to the DL \cap , \cup , and \neg
 $Q': (\text{PersonalName} \cap \text{TEST-C}(\text{rtrunc}_A)) \cap \text{Date}(1821) \text{ !!!!!}$

12

Z39.50 Query Translation to DL (cont)

- **Problem:** All the objects returned as an answer to a query must belong to a specific concept characterizing the source called **Central Concept** (Museum Object).
- **Idea:** Introduce Concept Path Expressions for each AP to connect through roles the individuals of the AP concept with the individuals of the Central Concept
- **Example:**

$E_{Date} = \exists \text{hasEvent.}(\exists \text{happenedIn.} \text{Time_Span})$

$E_{PersonalName} = (\exists \text{hasEvent.}(\exists \text{hasActor.} \text{Actor})) \cup (\exists \text{ownedBy.} \text{Owner})$

1
3

The 3 steps of Query Translation

Q: (PersonalName="A", Truncation="Right") AND Date=1821

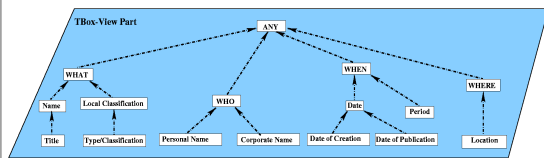
- Translate it into elementary DL query concepts
- Substitute the APs with their definitions using only primitive concepts
- Use the path expressions to consider only objects of the central concept and get the final answer:

$C_{Answer} = \{a \in O_{\Sigma} \mid \Sigma \models (\text{MuseumObject} \cap ((\text{OwnedBy.}(\text{Owner} \cap \text{TEST-C}(\text{rtrunc}_A))) \cup (\exists \text{hasEvent.}(\exists \text{hasActor.}(\text{Actor} \cap \text{TEST-C}(\text{rtrunc}_A)))))) \cap (\exists \text{hasEvent.}(\exists \text{happenedIn.} \text{Time_Span}(1821))) \} (a)\}$

1
4

Benefits of DL in AP Conceptualization

- Conceptual structuring of flat Z39.50 lists of Access Points using **DL Subsumption** reasoning service
 - Third party organizations: better AP design
 - End users: Better query formulation
 - Wrapper: Optimization of the query answering process



5

DL Benefits in Intelligent Query Processing (I)

- **Intentional query optimization:** Check if a query has contradictory description without accessing the actual data, e.g.:

Q₁: PersonalName="Androutsos" AND ProtectionStatus="Preserved"

- **Semantic query optimization,** e.g.:

Q₂: PersonalName="Androutsos" OR Who="Androutsos"

can be rewritten as:

Q₂': Who="Androutsos"

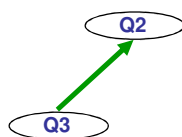
1
6

DL Benefits in Intelligent Query Processing (II)

- **Intelligent caching of the query results** given the stateful nature of the protocol, e.g.:

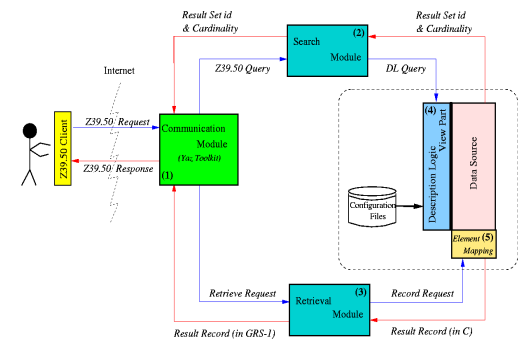
Q₃: Q₂ AND Date=1821

The caching results of Q₂ would contain only its proper individuals, not those that belong to Q₃



1
7

The Z39.50-DL Wrapper Toolkit Architecture



1
8

Summary and Future Work

- We addressed the declarative specification of Z39.50 wrappers
- We proposed a wrapper generator toolkit for expressing the mappings of APs to the source data based on DL
 - Can easily express many translation cases
 - Allows the formal validation of the quality of the mappings
 - Enables reasoning about the relationships between these mappings
 - Can be used for query answering, opening interesting opportunities for optimization
- **Current Status** of the toolkit:
 - Support of only the DL instance checking
- **Future Work:**
 - Complete the implementation of the remaining reasoning services
 - Experiment with other repositories technolog. (RDBMS, ODBMS, ...) ¹₉