



X3ML Mapping Framework for Information Integration

M. Theodoridou, N. Minadakis
FORTH - Institute of Computer Science, Greece

Seminarium om beständiga identifierare (persistent identifiers – PID)

March 7, 2016

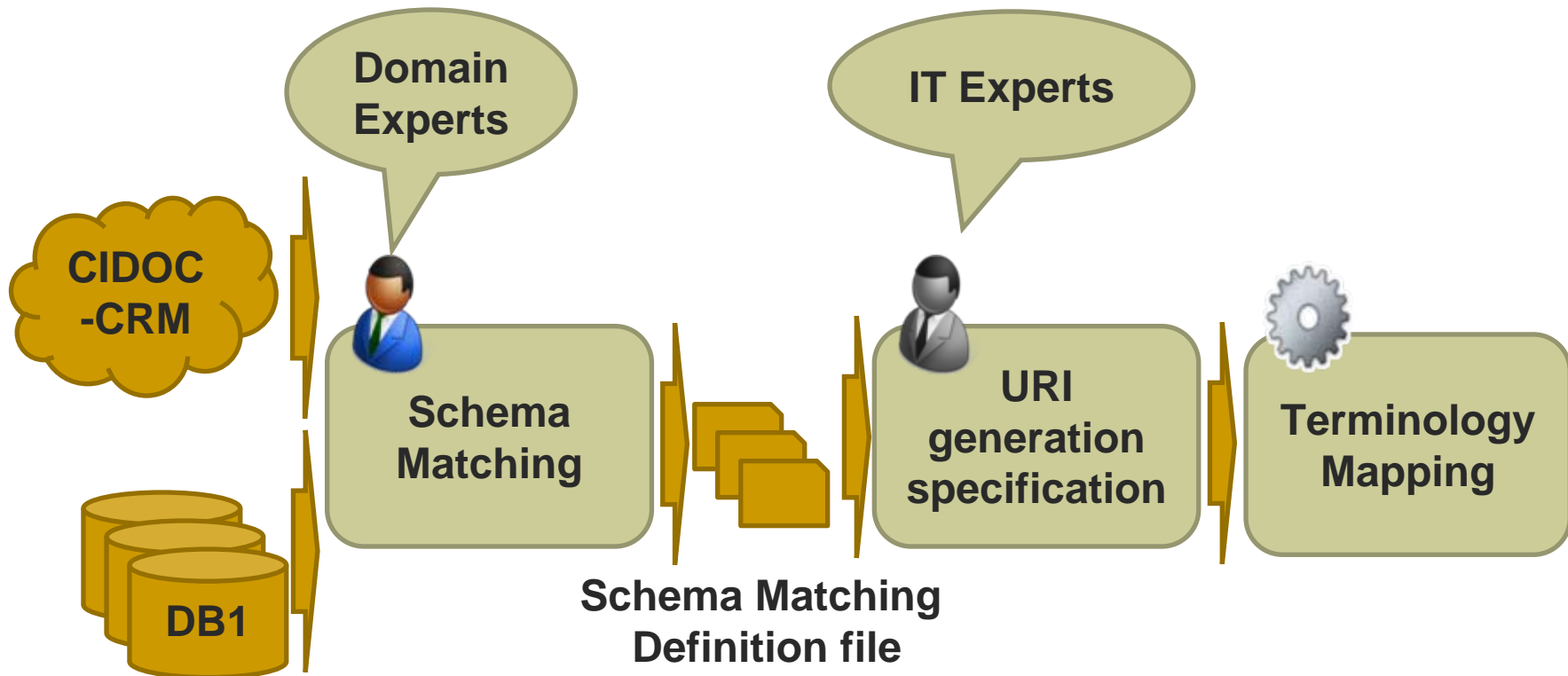


Synergy Reference Model

- A **reference model** for a better practice of **data provisioning** and **aggregation** processes.
- An initiative of the CIDOC CRM Special Interest Group.
- It is based on experience and evaluation of national and international information **integration projects**.
- It defines a consistent set of **business processes, user roles, generic software components** and **open interfaces** that form a harmonious whole



Transformation workflow





X3ML - URI generation policy

- The definition of the URI generation policy is a separate step and ***follows the schema matching.***
- It is performed usually by an ***IT expert*** who must ensure that the generated URIs match certain criteria such as consistency uniqueness.
- A set of ***predefined URI generators*** (UUIDs, literals) and ***templates*** are available but any URI generating function can be implemented and incorporated in the system.
- In the X3ML definition, the target domain and all range entities must contain functions that will generate ***URIs or literals.***

The result of the schema matching and URI generation policy steps is a ***complete X3ML mapping definition file*** that will be fed to the X3ML engine for the transformation of the data.



X3ML - URI generation policy

- Create a generator policy file (XML file)

- Generators through templates
 - Define patterns

- Custom Instance Generators
 - any URI generating function can be implemented and incorporated in the system
 - UUID
 - URIorUUID
 - Dates



sample URI generation policy file

```
<generator_policy>
<generator name="MyGeneratorFunction" prefix="MyNamespace">
  <pattern> {argument1}/{argument2} </pattern>
</generator>

<generator name="URLorUUID">
  <custom generatorClass="gr.forth.URLorUUID">
    <set-arg name="text"/>
  </custom>
</generator>
</generator_policy>
```



sample URI generation policy file

MyGeneratorFunction: PIDgenerator

MyNamespace: cb defined as <http://www.culturebroker.se/>

```
<generator name="PIDgenerator" prefix="cb">  
  <pattern> {type}/{id} </pattern>  
</generator>
```

```
<generator name="URLorUUID">  
  <custom generatorClass="gr.forth.URLorUUID">  
    <set-arg name="text"/>  
  </custom>  
</generator>
```



sample input record

<records>

<record>

<PID>**CB12345**</PID>

<otherID>**12**</otherID>

</record>

<record>

<PID>**CB56789**</PID>

<otherID>

<http://www.cb.se/Object/15>

</otherID>

</record>

</records>

P	↓ PID	↓ P1_is_identified_by	
		<input type="checkbox"/> E42_Identifier	
		Instance Generator Name	
		PIDgenerator	
		Argument	
		Name	Type
		type	constant
		Value	
		PID	
		Argument	
		Name	Type
		id	xpath
		Value	
		text()	
		Add Argument	
		Add Label Generator	



sample target record

```
<rdf:RDF
```

```
....
```

```
<crm:E22_Man-Made_Object rdf:about="http://www.cb.se/Object/15">
```

```
<crm:P1_is_identified_by>
```

```
<crm:E42_Identifier rdf:about="http://www.culturebroker.se/PID/CB56789"/>
```

```
</crm:P1_is_identified_by>
```

```
</crm:E22_Man-Made_Object>
```

```
<crm:E22_Man-Made_Object rdf:about="urn:uuid:72d7d450-1b84-4f96-a171-91bc73be6023">
```

```
<crm:P1_is_identified_by>
```

```
<crm:E42_Identifier rdf:about="http://www.culturebroker.se/PID/CB12345"/>
```

```
</crm:P1_is_identified_by>
```

```
</crm:E22_Man-Made_Object>
```

```
</rdf:RDF>
```



Mapping Memory Manager (3M)

<http://www.ics.forth.gr/isl/3M/>

URI example: Mapping 516

Maria Theodoridou, Nikos Minadakis
FORTH - Institute of Computer Science, Greece
{maria,minadakn}@ics.forth.gr



Thank you!