

DataMaster

a Plug-in for importing schema and data from
databases into Protégé

Csongor Nyulas

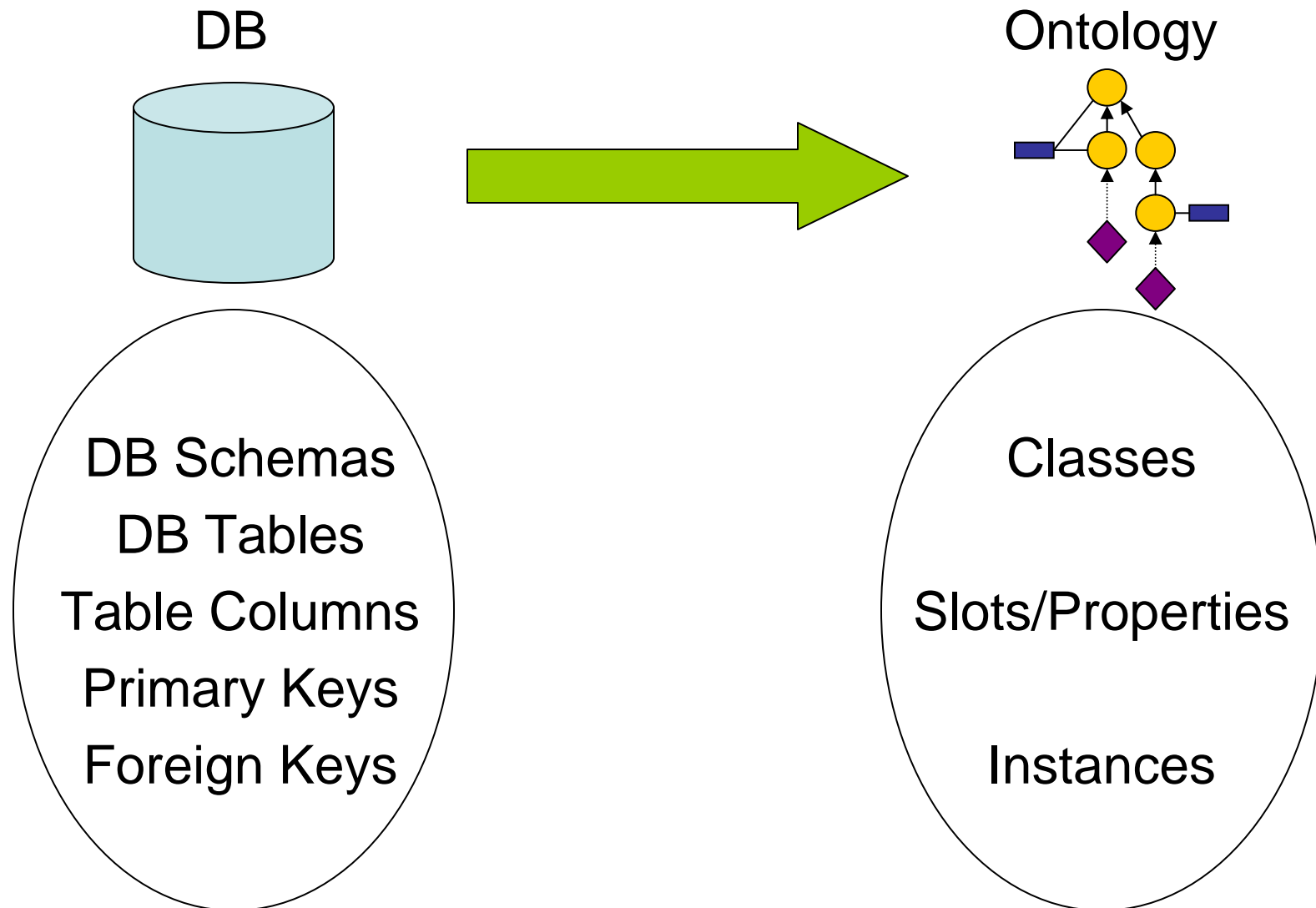
Martin O'Connor, Samson Tu

Stanford University - SMI

Background and Motivation

- The BioSTORM project
- Integration of multiple data sources
- Protégé plug-in available to the user community
- Use cases:
 - Ontology to Relational DB mapping
 - Import of DB content

Functionality



The DataMaster Plug-in

SurveillanceMethod Protégé 3.3 beta (file:\C:\work\BioSTORM\examples\SurveillanceMethod20070306\SurveillanceMethod.pprj, OWL / RDF Files)

File Edit Project OWL Code Tools Window Help

Metadata (SurveillanceMethods.owl) OWLClasses Properties Individuals Forms DataMaster v1.0

Data Source Type ☐ ODBC ☒ JDBC

JDBC Driver com.mysql.jdbc.Driver

JDBC URL jdbc:mysql://localhost:3306/quebec

User Login csongor

Password *****

Connect Info

Select superclass for the table classes:

- rdfs:Class
- rdf:List
- dd:DataDescription
 - dd:DataGroup
 - ds:QuebecDataCase
 - ds:_911CallRecord
 - ds:NormalizedDataGroup
 - ds:QuebecDataOriginalView**
 - ds:VariableDataGroup
- dd:DataType

Import location:

- ☒ in the current ontology
 - ☒ use different namespaces
- ☐ in a separate ontology

Import tables as: ☒ classes ☐ Relational OWL instances

☐ Include table name in column name

Define DB column types by:

- ☒ hasColumnType property referring to ColumnType instances
- ☐ hasXSDType property
- ☐ rdfs:range property (as in Relational OWL)

☐ Import table content

Data Tables

☐ System Table ☒ Table ☒ View ☐ Other

- claims
- drugs
- people

☒ Select all tables

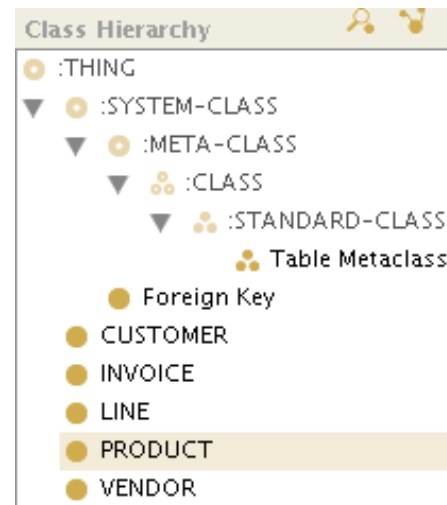
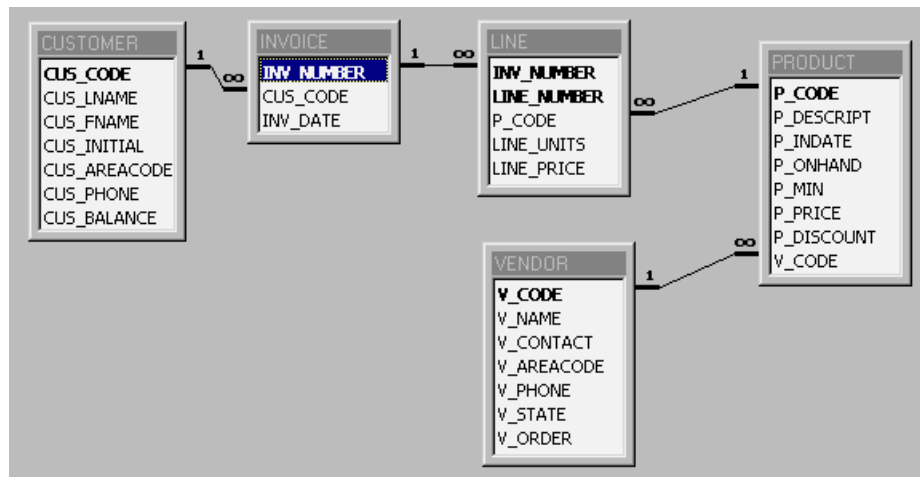
☒ Preview claims

Number of rows 50 (<0 = all) Sgt

PID	CLPROF	NOPROF	DIAG	ETABL	ETABU	CACTE	RACTE	DT	PAYE	CLREF	NOREF	REMUNE
GEFL44381...		1 35R17	4869	0X7	O	51	1	2002026	18,000		1 36109	G
ICCU0M360...		1 35R17	4869	0X7	O	51	1	2002026	18,000		1 36109	G
YVIE46480...		1 35R17	4869	0X3	O	41	1	2002027	2,533		0 00000	G
NWIE25940...		1 35R17	4869	0X3	O	41	1	2002027	2,533		0 00000	G
GNXE45230...		1 35R17	4869	0X3	O	39	1	2002027	1,900		0 00000	G
DEVA4345...		1 35R17	4869	0X3	O	39	1	2002027	1,900		0 00000	G
ICCU0M360...		1 2HHR8	4869	0X3	O	39	1	2002028	1,900		0 00000	G
PIEB38850551		1 2HHR8	4869	0X3	O	39	1	2002028	1,900		0 00000	G
EGXF200R0		1 2HHR8	4869	0X3	O	34	1	2002028	5,400		0 00000	G

Import

Import in Protégé Frames



Template Slots

Name	Cardinality	Type
P_CODE	single	Integer
P_DESCRIPT	single	String
P_DISCOUNT	single	Float
P_INDATE	single	String
P_MIN	single	Integer
P_ONHAND	single	Boolean
P_PRICE	single	Float
V_CODE	single	Integer

Primary Key Fields

P_CODE

Foreign Keys

◆ Foreign_Key_Instance_4

Import in Protégé OWL

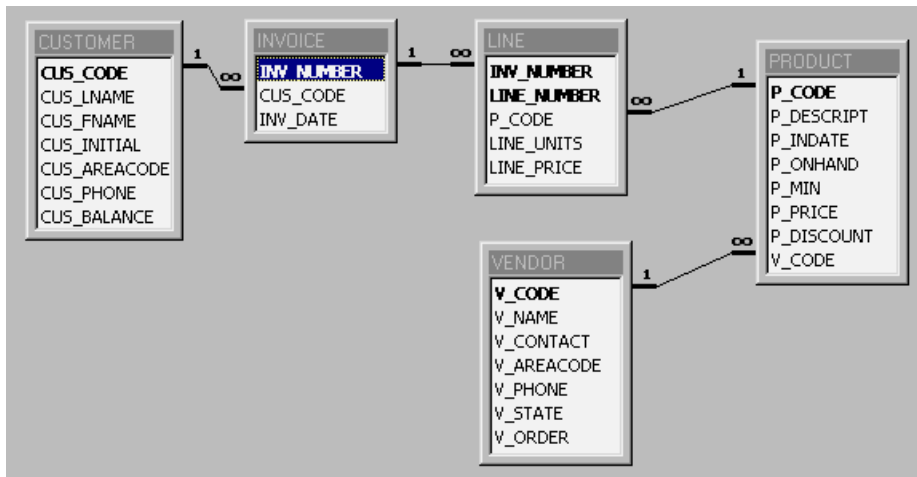
- 3 options for representing DB Tables
 - As OWL classes
 - As instances of Relational.OWL classes
 - As OWL classes, instances of Relational.OWL classes

Import in Protégé OWL

- Represent Tables as OWL Classes -

Namespace Prefixes

Prefix	Namespace
xsd	http://www.w3.org/2001/XMLSchema#
rdfs	http://www.w3.org/2000/01/rdf-schema#
rdf	http://www.w3.org/1999/02/22-rdf-syntax-ns#
owl	http://www.w3.org/2002/07/owl#
db	http://www.dbs.cs.uni-duesseldorf.de/RDF/relational.owl#
db	http://biostorm.stanford.edu/db_table_classes?DSN=jdbc:mysql://localhost:3306/trade#



Asserted Hierarchy

- owl:Thing
 - db:ForeignKey
 - db:CUSTOMER
 - db:INVOICE
 - db:LINE
 - db:PRODUCT
 - db:VENDOR

For Class: db:PRODUCT

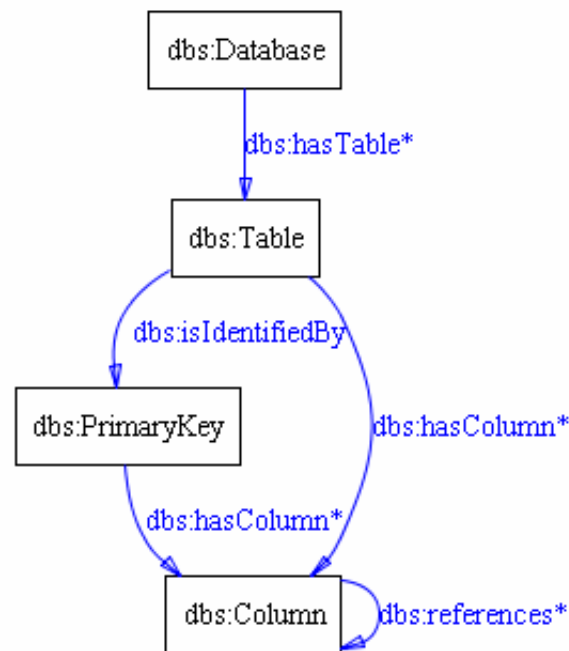
Property	
rdfs:comment	
db:hasForeignKeys	FK: PRODUC
db:hasPrimaryKeyFields	P_CODE
db:isBridgeTable	false

- db:PRODUCT.P_CODE (single int)
- db:PRODUCT.P_DESCRIPT (single string)
- db:PRODUCT.P_DISCOUNT (single float)
- db:PRODUCT.P_INDATE (single date)
- db:PRODUCT.P_MIN (single int)
- db:PRODUCT.P_ONHAND (single boolean)
- db:PRODUCT.P_PRICE (single float)
- db:PRODUCT.V_CODE (single int)

Import in Protégé OWL

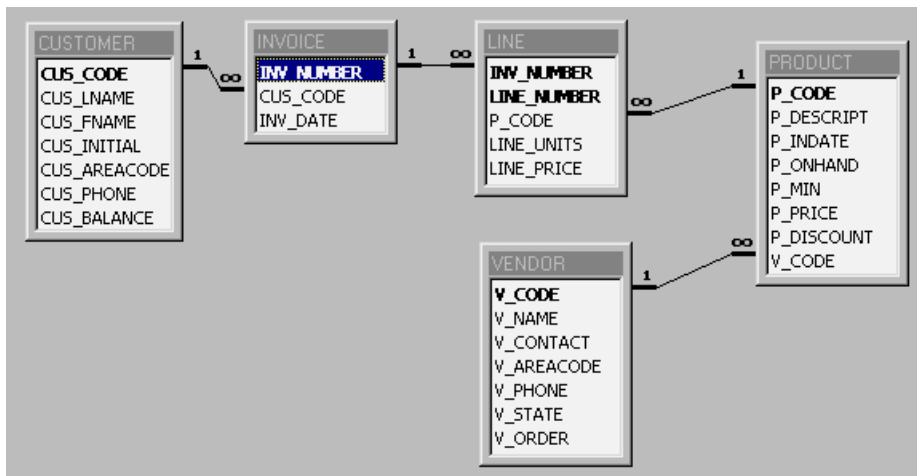
- using the Relational.OWL ontology -

- The Relational.OWL ontology:



Import in Protégé OWL

- Represent Tables as instances of Relational.OWL classes -



Class Hierarchy

- owl:Thing
 - db:Database (1)
 - db:Table (5)
 - db:Column (29)
 - db:PrimaryKey (5)
 - db:ColumnType (5)

For Class: db:Table

Asserted **Inferred**

Asserted Instances

- db:CUSTOMER
- db:INVOICE
- db:LINE
- db:PRODUCT
- db:VENDOR

db:isIdentifiedBy

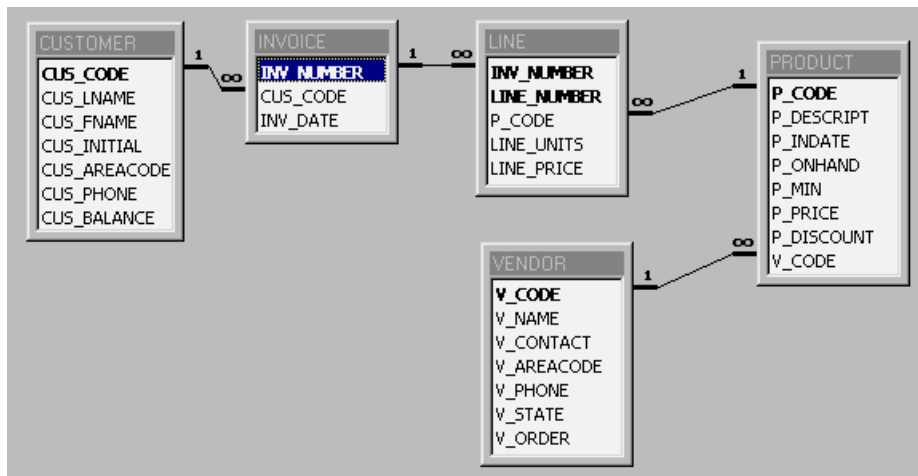
- db:PK_PRODUCT

db:hasColumn

- db:PRODUCT.P_CODE
- db:PRODUCT.P_DESCRIPT
- db:PRODUCT.P_INDATE
- db:PRODUCT.P_ONHAND
- db:PRODUCT.P_MIN
- db:PRODUCT.P_PRICE

Import in Protégé OWL

- Represent Tables as classes AND Relational.OWL instances -



Class Hierarchy

- owl:Thing
 - db1:Database (1)
 - db1:Table (5)
 - db1:Column (29)
 - db1:PrimaryKey (5)
 - db1:ColumnType (5)
 - db1:CUSTOMER
 - db1:INVOICE
 - db1:LINE
 - db1:PRODUCT
 - db1:VENDOR

Asserted Instances

- db1:CUSTOMER
- db1:INVOICE
- db1:LINE
- db1:PRODUCT
- db1:VENDOR

db1:isIdentifiedBy

db1:PK_PRODUCT

db1:hasColumn

- db1:PRODUCT.P_CODE
- db1:PRODUCT.P_DESCRIPT
- db1:PRODUCT.P_INDATE
- db1:PRODUCT.P_ONHAND
- db1:PRODUCT.P_MIN

DB Concept Representations

<div> <div>Ontology</div> <div>Database</div> </div>	Protégé Frames	Protégé OWL		
		Type I	Type II	Type II
DB Schema	-	Implicitly by namespace	<i>Database</i> class inst.	<i>Database</i> class inst.
DB Table	Class	Class	<i>Table</i> class instances	<i>Table</i> meta-class inst.
Table Column	Template slots	Datatype Properties	<i>Column</i> class inst.	Inst. <i>Column</i> meta-prop.
Primary Key	<i>Primary Key Fields</i> slot	-	<i>PrimaryKey</i> class inst.	<i>PrimaryKey</i> class inst.
Foreign Key	<i>Foreign Key</i> instances	<i>ForeignKey</i> instances	<i>references</i> property	<i>references</i> property

Conclusions

- DataMaster: A plug-in available as part of the Protégé distribution
- Works with Protégé Frames and Protégé OWL
- Imports schema (and data) from DB
- Multiple import options
- Intuitive user interface