

# Rdfization of a relational database from medicine domain

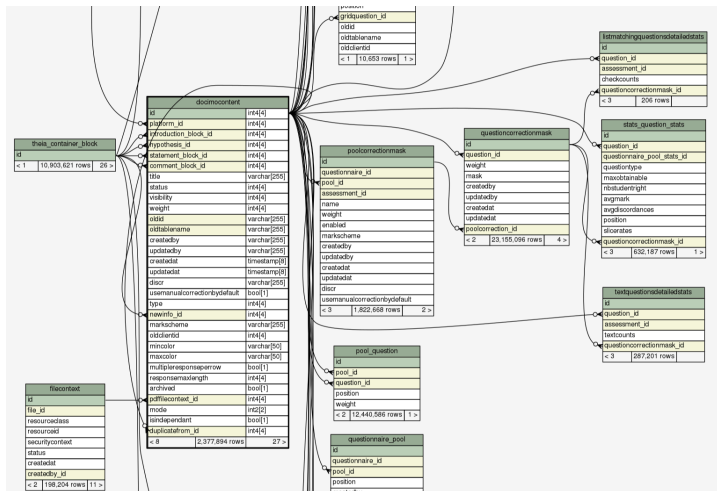
Olivier Palombi, Marie-Christine Rousset, Fabrice Joanot,  
Adam Sanchez

July 3, 2018

SIDES is a learning digital platform common to all French medical schools, used for official exams (tests) in faculties and for the training of students for the National Ranking Exam (ECN) which is fully computerized since 2016 (ECNi). SIDES allows to trainers:

- create test exams
- perform an automatic correction of exams
- analyze the relevance and discrimination of tests

# The schema of the relational database of SIDES



# The schema of the relational database of SIDES

- The number of tables has increased from 126 in 2015 to 357 in 2018.
- The mediation of IT experts is necessary to build complex SQL queries.

# The challenges

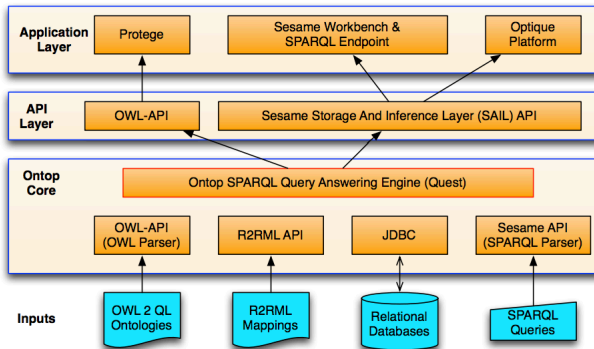
- How to adapt the schema of the database when changes in the underlying knowledge are fast and/or composite.
- How to get more knowledge from the existing one.
- How to state more information for relationships between tables (foreign keys).

# The challenges

- How to adapt the schema of the database when changes in the underlying knowledge are fast and/or composite.
- How to get more knowledge from the existing one.
- How to state more information for relationships between tables (foreign keys).

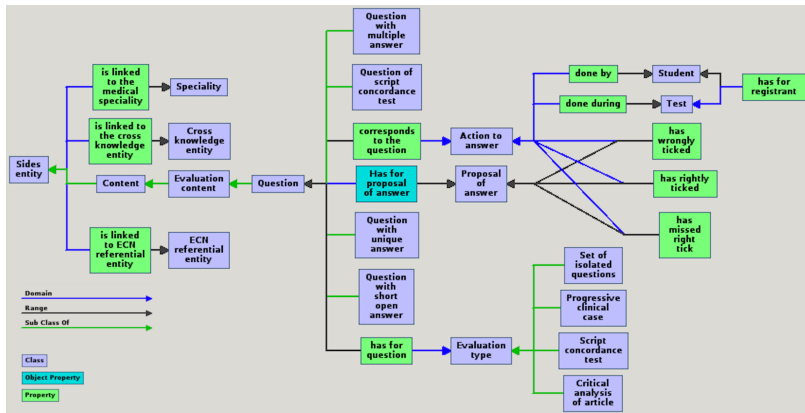
The solution to tackle these challenges was the implementation of an OBDA framework.

# Ontop: a framework for OBDA



It enables querying virtual RDF graphs using SPARQL by translating the queries into SQL. [1]

# Input 1: The Ontosides ontology





# The Ontosides ontology

	Total
Number of Classes	67
Number of Properties	39

Table: Number of classes and properties

# Input 2: The Ontop mappings

```
[PrefixDeclaration]
: http://www.side-sante.fr/sides#
sp: http://spinrdf.org/sp#
owl: http://www.w3.org/2002/07/owl#
rdf: http://www.w3.org/1999/02/22-rdf-syntax-ns#
spl: http://spinrdf.org/spl#
xml: http://www.w3.org/XML/1998/namespace
xsd: http://www.w3.org/2001/XMLSchema#
rdfs: http://www.w3.org/2000/01/rdf-schema#
spin: http://spinrdf.org/spin#
sides: http://www.side-sante.fr/sides#
wiki_sides: http://wiki.side-sante.fr/doku.php?id=sides:ref:

[SourceDeclaration]
sourceUri datasource1
connectionUrl jdbc:postgresql://22.43.56.78:5432/sides
username postgres
password y73hnfnjnrrwww
driverClass org.postgresql.Driver

[MappingDeclaration] @collection []

mappingId urn:test
target sides:test{id} a sides:test .
source select a.id as id, a.title, a.startdate, a.enddate from public.assessment a

mappingId urn:test_has_for_title
target sides:test{id} sides:has_for_title "{title}"^^xsd:string .
source select a.id as id, a.title as title, a.startdate, a.enddate from public.assessment a

]]
```



# The Ontop mappings

The screenshot displays the Ontop Mapping Assistant interface. The main window is titled "Mapping editor" and contains a "Mapping manager" pane on the left and an "Edit Mapping" dialog box in the foreground.

**Mapping manager (Left Pane):**

- urn:test**  
sides:test[id] a sides:test .  
select a.id as id, a.title, a.startdate, a...
- urn:test\_has\_for\_title** (highlighted)  
sides:test[id] sides:has\_for\_title [title]^xsd:string  
select a.id as id, a.title as title, a.sta...
- urn:test\_start\_and\_end\_dates**  
sides:test[id] sides:starting\_date\_of\_test {startda...  
select a.id as id, a.title, a.startdate as...
- urn:evaluation\_has\_for\_textual\_content**  
sides:eval[id] sides:has\_for\_textual\_content {eval...  
SELECT di.id as id, string\_agg(di.introduc...  
inner join ontosides.evaluation ev...  
group by di.id
- urn:relation\_evaluation\_type\_question**  
sides:eval[pool\_id] sides:has\_for\_question sides:c...  
select pg.pool\_id as pool\_id, pg.question\_...  
inner join ontosides.evaluation ev...
- urn:relation\_test\_evaluation\_type**  
sides:test[assessment\_id] sides:is\_made\_of sides:...  
select dca.assessment\_id as assessment\_id,...  
inner join ontosides.evaluation ev...
- urn:relation\_test\_student**  
sides:test[assessment\_id] sides:has\_for\_registran...  
select p.assessment\_id as assessment\_id, p.participant\_id as participant\_id from public.participant p

**Edit Mapping Dialog (Right Pane):**

- Mapping ID:** urn:test\_has\_for\_title
- Target (Triples Template):** sides:test[id] sides:has\_for\_title [title]^xsd:string .
- Source (SQL Query):** select a.id as id, a.title as title from public.assessment a limit 10
- Test SQL Query:** A table with columns 'id' and 'title'.

id	title
1	CONF-2015-2016-HEMATO n*1- M2-Gpe F-DM
13894	PHARMA 3A EVAL UE 2014-2015 - 3A S05 - 1er déce...
11	C6-ORL-02/09/2015-DP1
15567	SDP
16	CONFBDI/CHIRVISC/DFASM3/01
17	CONFBDI/MAL INF/DFASM3/01
- Buttons:** Update (green checkmark), Cancel (red X)

**Bottom Bar:**

- Mapping count: 189
- Search: pred:action\_to\_answer
- Enable filter:
- To use the reasoner click Reasoner > Start reasoner  Show Inferences

# The Ontop query editor

The screenshot shows the Ontop query editor interface. At the top, there are tabs for 'Query', 'Ontop Mappings', 'Ontop SPARQL', and 'VOWL'. The main area contains a SPARQL query:

```
Query Editor
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX sides: <http://www.side-sante.fr/sides#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

select * {
  ?s rdf:type sides:question.
}
limit 10
```

Below the query, the execution status is shown: 'Execution time: 0.275 sec - Number of rows retrieved: 10'. There are controls for 'Show: 100', 'All', 'Short IRI', 'Attach Prefixes', 'Execute', and 'Save Changes'. The results pane shows a list of 10 URIs:

```
s
<http://www.side-sante.fr/sides#q28318>
<http://www.side-sante.fr/sides#q28341>
<http://www.side-sante.fr/sides#q28343>
<http://www.side-sante.fr/sides#q28344>
<http://www.side-sante.fr/sides#q28346>
<http://www.side-sante.fr/sides#q28348>
<http://www.side-sante.fr/sides#q28349>
<http://www.side-sante.fr/sides#q28350>
<http://www.side-sante.fr/sides#q28355>
<http://www.side-sante.fr/sides#q28356>
```

# A workaround for RDF containers

```
mappingId urn:action_to_answer
target   sides:adr{response_id_new} a sides:action_to_answer ; sides:has_for_timestamp "{(begindate)}"^^xsd:dateTime .
source   select response_id_new, begindate from ontosides.response

mappingId urn:relation_action_to_answer_test
target   sides:adr{response_id_new} sides:done_during sides:test{assessment_id} .
source   select response_id_new, assessment_id from ontosides.response

mappingId urn:relation_evaluation_list_of_questions
target   sides:lq{pool_id} a rdf:Seq . sides:eval{pool_id} sides:has_for_list_of_questions sides:lq{pool_id} .
source   select ev.pool_id as pool_id from ontosides.evaluation ev inner join
         public.docimocontent d on d.id = ev.pool_id where d.discr = 'dp'

mappingId urn:sequence_of_questions_by_evaluation
target   sides:lq{pool_id} |rdf:_{pos}| sides:q{question_id} .
source   select pq.pool_id as pool_id, pq.question_id as question_id, (pq.position + 1) as pos from public.pool_question pq
         inner join ontosides.evaluation ev on ev.pool_id = pq.pool_id
```

# A workaround for RDF containers

```
@prefix sides: <http://www.side-sante.fr/sides#> .
@prefix rdf:   <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .

sides:eval830 sides:has_for_list_of_questions sides:lq830 .
sides:lq830   rdf:type rdf:Seq ;
  rdf:_1     sides:q330661 ;
  rdf:_2     sides:q330662 ;
  rdf:_3     sides:q330663 ;
  rdf:_4     sides:q330664 ;
  rdf:_5     sides:q330665 ;
  rdf:_6     sides:q330666 ;
  rdf:_7     sides:q330667 ;
  rdf:_8     sides:q330668 ;
  rdf:_9     sides:q330669 ;
  rdf:_10    sides:q330670 ;
  rdf:_11    sides:q330671 ;
  rdf:_12    sides:q330672 ;
  rdf:_13    sides:q330673 ;
  rdf:_14    sides:q330674 ;
  rdf:_15    sides:q330675 .
```

# A workaround for RDF containers

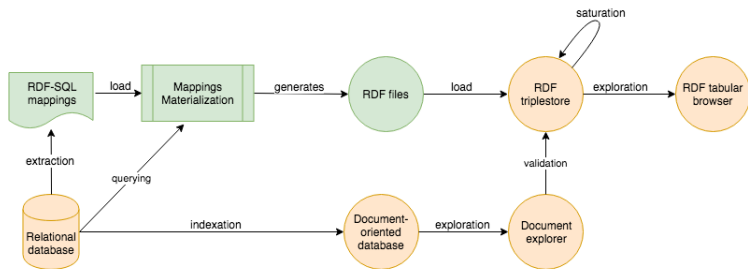
```
@prefix sides: <http://www.side-sante.fr/sides#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .

sides:eval830 sides:has_for_list_of_questions sides:lq830 .
sides:lq830 rdf:type rdf:Seq ;
  rdf:_1 sides:q330661 ;
  rdf:_2 sides:q330662 ;
  rdf:_3 sides:q330663 ;
  rdf:_4 sides:q330664 ;
  rdf:_5 sides:q330665 ;
  rdf:_6 sides:q330666 ;
  rdf:_7 sides:q330667 ;
  rdf:_8 sides:q330668 ;
  rdf:_9 sides:q330669 ;
  rdf:_10 sides:q330670 ;
  rdf:_11 sides:q330671 ;
  rdf:_12 sides:q330672 ;
  rdf:_13 sides:q330673 ;
  rdf:_14 sides:q330674 ;
  rdf:_15 sides:q330675 .
```

The solution was to extend the class `ModellOManager` to generate this kind of mapping.



# The workflow to materialize Ontop mappings



# The RDF data of Ontosides

- 1 475 080 485 RDF triples.
- 71 257 evaluations
- 590 654 questions
- 2 835 718 proposal of answer
- 276 555 293 action to answer
- 64 957 students

- Implementation of SHACL tests for RDF data validation.
- Implementation of rules to generate more RDF data.

Thanks!

- [1] D. Calvanese, B. Cogrel, E. G. Kalayci, S. Komla-Ebri, R. Kontchakov, D. Lanti, M. Rezk, M. Rodriguez-Muro, and G. Xiao.  
OBDA with the ontop framework.  
*In 23rd Italian Symposium on Advanced Database Systems, SEBD 2015, Gaeta, Italy, June 14-17, 2015.*, pages 296–303, 2015.