

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/352816605>

VIVO-STUDIO: TOWARDS A SOFTWARE THAT FACILITATES AND STANDARDIZES THE PROCESS OF DATA TRANSFORMATION FROM INSTITUTIONAL SOURCES TO VIVO FOR MULTIDISCIPLINARY TEAMS 12th Annual VIVO...

Presentation · June 2021

DOI: 10.13140/RG.2.2.16343.60328/1

CITATIONS

0

READS

74

1 author:



Michel Héon

Université du Québec à Montréal

49 PUBLICATIONS **157** CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Ontology CASE tool for Graphical Web Ontology Language (OntoCASE4GOWL) [View project](#)



Semantic deepLearning [View project](#)

VIVO-STUDIO: TOWARDS A SOFTWARE THAT FACILITATES AND STANDARDIZES THE PROCESS OF DATA TRANSFORMATION FROM INSTITUTIONAL SOURCES TO VIVO FOR MULTIDISCIPLINARY TEAMS

12th Annual VIVO Conference 2021

Michel Héon
Université du Québec à Montréal (UQAM)
heon.michel@uqam.ca

Abstract

An essential step in the production of VIVO's institutional instance is the data provisioning with relevant, coherent, up to date and valid data. Complex data transformation processes must be developed and implemented within the institution by several specialists coming from various fields of expertise. In most cases, the data transformation process can be divided into the following tasks : selecting the data to be extracted; extracting data from various institutional data sources; converting tabular data representation into knowledge graphs; editing ontologies and vocabularies in RDF/S; data mapping from a source vocabulary to VIVO's vocabulary; data ingestion in a local VIVO instance and a final step of verifying and validating the data managed by VIVO. The process involves the collaboration of several professionals who, on the one hand, design the data transformation rules specific to each institution, and on the other hand, implement the transformation rules in various software modules in order to automate the VIVO's data loading.

A critical issue that ensures the success of institutional data integration in VIVO is related to the quality of the collaboration between the various professionals involved, the simplicity of carrying out the recurring tasks (e.g.: migrating dataset from JSON notation to TURTLE notation, or executing SPARQL queries in a graph) and access of encapsulated and preconfigured services (e.g. a local and pre-installed VIVO instance). To facilitate collaboration, we have developed VIVO-Studio, a software tool with scalable, adaptive and incremental features that facilitates and standardize the tasks required in the data transformation process. VIVO-Studio is intended for computer scientists at all levels as well as for ontologists responsible for data quality, whether they are librarians, researchers, data scientists or database administrators. VIVO-Studio encapsulates a set of software tools such as a Tomcat server, access to Java APIs from Apache-Jena and OWLApi, an Apache Fuseki server, Apache Kafka services and many others. During our presentation, we will present the various contexts and use cases for VIVO-Studio. . Thereafter, we will present the component architecture of VIVO-Studio as well as the principal functionalities which compose it and which will be, as a demonstration guide, supported by some screen captures.

Why This Talk?

Introducing VIVO-Studio, the context of its use as well as the demonstration of some use cases

What Is VIVO-Studio ?

VIVO-Studio is an initiative of the *Université du Québec à Montréal* whose objective is to design a software assistant to facilitate the software and ontological maintenance of VIVO which is accessible to teams of multidisciplinary professionals (software engineer, librarians, ontologists, data scientists, etc.)

Software Context

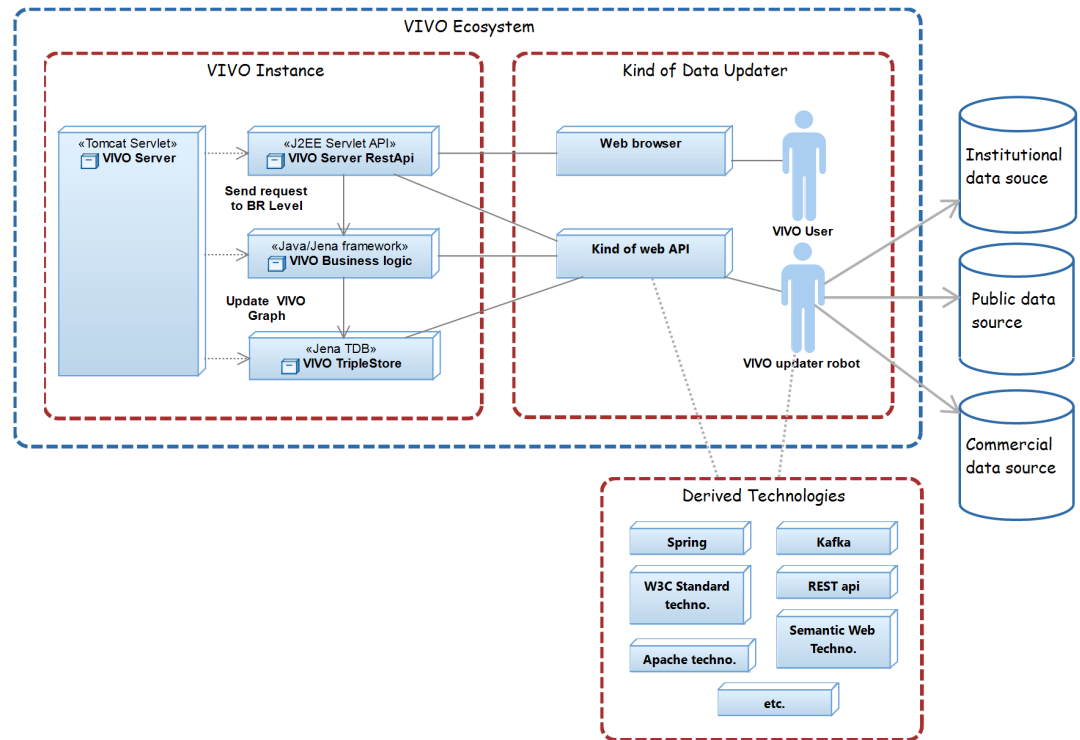
VIVO is designed on a 3-tier classic architecture:

- User interface
- Business logic layer
- Graphical data storage layer

Data in VIVO are updated

- By humans via a web browser
- By robots which are fed by data sources of heterogeneous structure (Relational DB, Web Data Source, Graphical DB) which are either: institutional, public or commercial

The robot exploits a multitude of derived technologies to meet specific needs

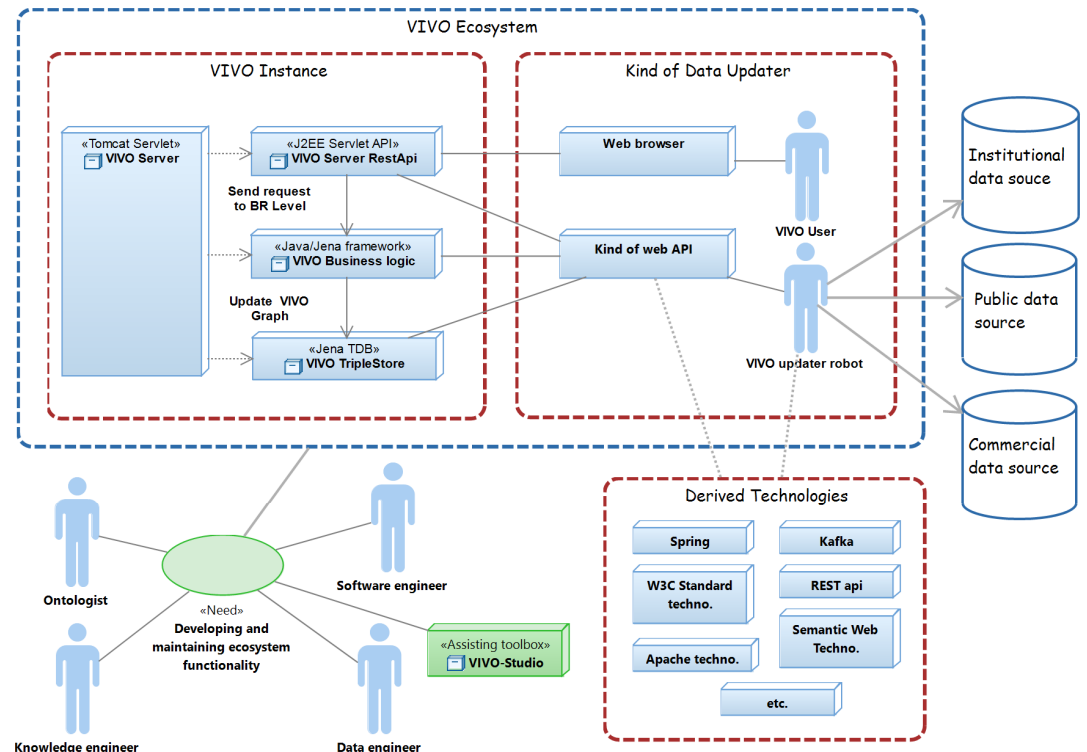


VIVO-Studio In The VIVO Ecosysem

VIVO-Studio is an assistance tools for the development and maintenance of VIVO features

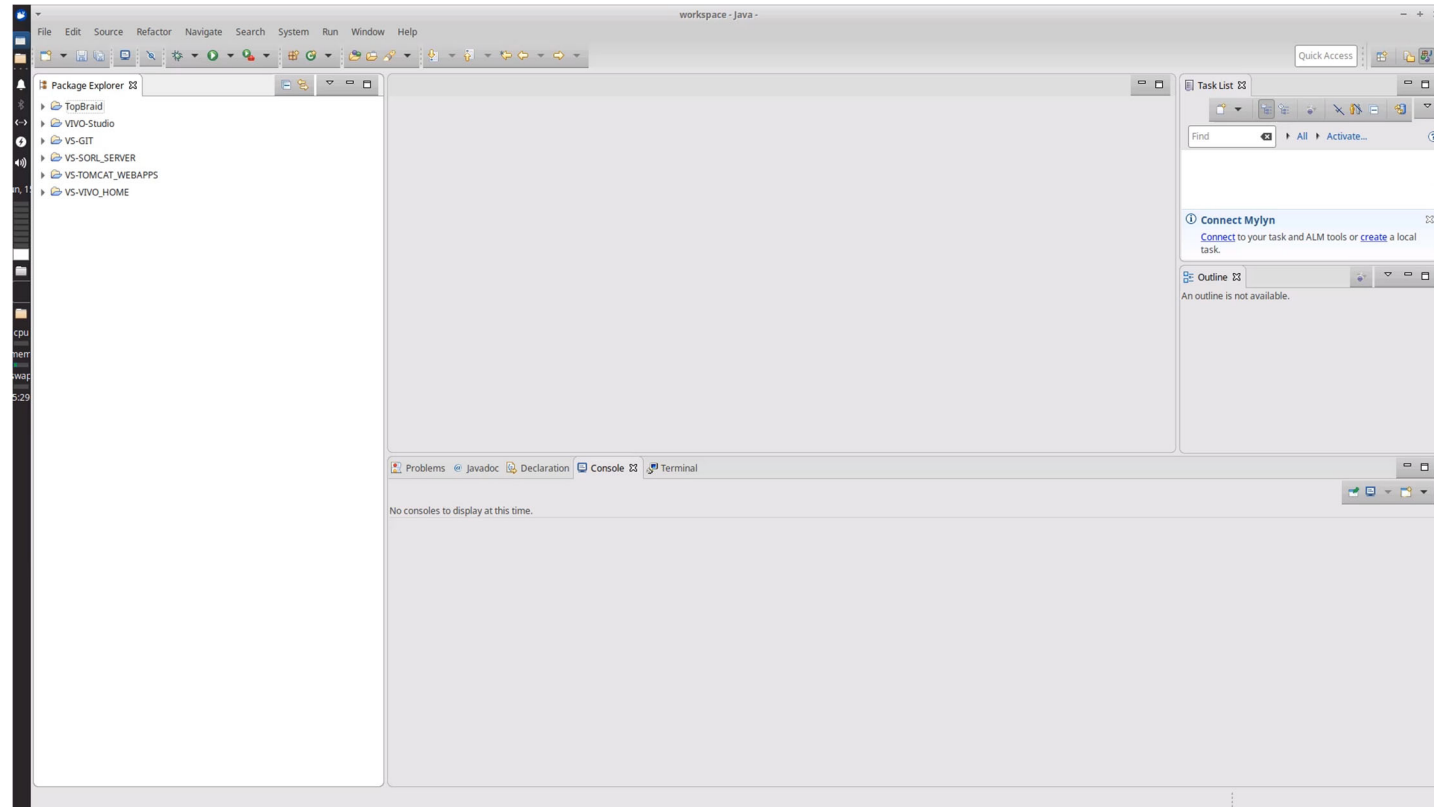
VIVO-Studio is intended for computer scientists at all levels as well as for ontologists responsible for data quality, whether they are librarians, researchers, data scientists or database administrators.

VIVO-Studio encapsulates a set of software tools such as a Tomcat server, access to Java APIs from Apache-Jena and OWLapi, an Apache Fuseki server, Apache Kafka services and many others.



DEMONSTRATION

Vivo-Studio Projects Organization



<https://youtu.be/Xu7YTe67H7w>

Installing and run VIVO

- a) Opening a terminal
- b) Clone VIVO from GIT
 - > vivo_clone.sh
- c) Compile VIVO
 - > vivo_compile.sh
- d) Start SOLR & Tomcat
- e) Start using VIVO

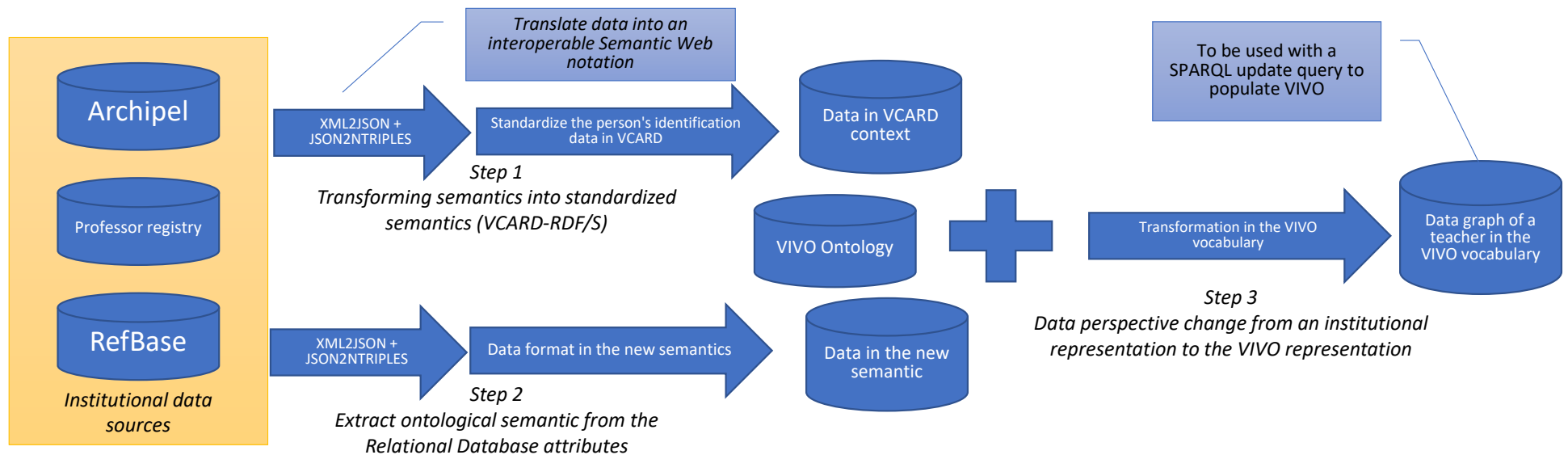
https://youtu.be/uq2bRTTTr_U

UseCase-2 Data transformation

The case of "adding a professor"

Use Vivo-Studio to transform data from institutional data sources to the VIVO TripleStore

The transformation workflow



Workflow realisation in VIVO-Studio

- a) Experimental context presentation
- b) On the fly file transformation JSON to Turle
- c) Executing TheData Transformation WorkFlow
- d) Validate the transformation with VIVO

<https://youtu.be/7xX4ps9w9fg>

Complexity of SPARQL update queries

```
prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix vcard: <http://www.w3.org/2006/vcard/ns#>
construct {
  ?uuid <http://uqam.ca/ontology#hasUUID> ?key .
  ?uuid rdf:type <http://purl.obolibrary.org/obo/BFO_0000001> .
  ?uuid <http://vitro.mannlib.cornell.edu/ns/vitro/0.7#mostSpecificType> <http://vivoweb.org/ontology/core#FacultyMember> .
  ?uuid rdf:type <http://xmlns.com/foaf/0.1/Agent> .
  ?uuid rdf:type <http://xmlns.com/foaf/0.1/Person> .
  ?uuid rdf:type <http://purl.obolibrary.org/obo/BFO_0000004> .
  ?uuid <http://www.w3.org/2000/01/rdf-schema#label> ?nomComple .
  ?uuid rdf:type <http://purl.obolibrary.org/obo/BFO_0000002> .
  ?uuid rdf:type <http://vivoweb.org/ontology/core#FacultyMember> .
  ?uuid rdf:type <http://www.w3.org/2002/07/owl#Thing> .
  ?uuid <http://purl.obolibrary.org/obo/ARG_2000028> ?n1 .
  ?n1 rdf:type <http://purl.obolibrary.org/obo/IAO_0000030> .
  ?n1 <http://purl.obolibrary.org/obo/ARG_2000029> ?uuid .
  ?n1 <http://vitro.mannlib.cornell.edu/ns/vitro/0.7#mostSpecificType> <http://www.w3.org/2006/vcard/ns#Individual> .
  ?n1 <http://www.w3.org/2006/vcard/ns#hasName> ?n2 .
  ?n1 rdf:type <http://purl.obolibrary.org/obo/ARG_2000379> .
  ?n1 rdf:type <http://www.w3.org/2002/07/owl#Thing> .
  ?n1 rdf:type <http://www.w3.org/2006/vcard/ns#Kind> .
  ?n1 rdf:type <http://purl.obolibrary.org/obo/BFO_0000031> .
  ?n1 rdf:type <http://purl.obolibrary.org/obo/BFO_0000002> .
  ?n1 rdf:type <http://purl.obolibrary.org/obo/BFO_0000001> .
  ?n1 rdf:type <http://www.w3.org/2006/vcard/ns#Individual> .
  ?n2 <http://www.w3.org/2006/vcard/ns#givenName> ?prenom .
  ?n2 <http://www.w3.org/2006/vcard/ns#familyName> ?nom .
  ?n2 <http://vitro.mannlib.cornell.edu/ns/vitro/0.7#mostSpecificType> <http://www.w3.org/2006/vcard/ns#Name> .
  ?n2 rdf:type <http://www.w3.org/2006/vcard/ns#Name> .
  ?n2 rdf:type <http://www.w3.org/2006/vcard/ns#Communication> .
  ?n2 rdf:type <http://www.w3.org/2006/vcard/ns#Explanatory> .
  ?n2 rdf:type <http://www.w3.org/2006/vcard/ns#Addressing> .
  ?n2 rdf:type <http://www.w3.org/2006/vcard/ns#Identification> .
  ?n2 rdf:type <http://www.w3.org/2002/07/owl#Thing> .
  ?n1 <http://www.w3.org/2006/vcard/ns#hasEmail> ?n3 .
  ?n3 rdf:type <http://www.w3.org/2002/07/owl#Thing> .
  ?n3 rdf:type <http://www.w3.org/2006/vcard/ns#Addressing> .
  ?n3 rdf:type <http://www.w3.org/2006/vcard/ns#Communication> .
  ?n3 rdf:type <http://www.w3.org/2006/vcard/ns#Explanatory> .
  ?n3 rdf:type <http://www.w3.org/2006/vcard/ns#Identification> .
  ?n3 rdf:type <http://www.w3.org/2006/vcard/ns#Work> .
  ?n3 rdf:type <http://www.w3.org/2006/vcard/ns#Email> .
  ?n3 <http://vitro.mannlib.cornell.edu/ns/vitro/0.7#mostSpecificType> <http://www.w3.org/2006/vcard/ns#Email> .
  ?n3 <http://vitro.mannlib.cornell.edu/ns/vitro/0.7#mostSpecificType> <http://www.w3.org/2006/vcard/ns#Work> .
  ?n3 <http://www.w3.org/2006/vcard/ns#email> ?email .
} where {
  bind("$NOM_COMPLET"@fr-ca as ?nomComple .
  bind("$NOM"@fr-ca as ?nom) .
  bind("$PRENOM"@fr-ca as ?prenom) .
  bind("<encas>($N)"@uqam.ca" as ?email) .
  bind(UUID() as ?n1) . bind(UUID() as ?n2) . bind(UUID() as ?n3) .
  ?uuid <http://uqam.ca/ontology#hasUUID> [] .
  ?uuid <http://uqam.ca/ontology#hasUUID> ?key .
}
```

VIVO-CONF. 2021 - TOWARD VIVO-STUDIO - MICHEL HÉON

Ontology modeling With TopbraidComposer

- a) Opening TopBraid Perspective
- b) Inside perspective views
- c) Visualize data using the
VIVO ontology vocabulary

<https://youtu.be/ye3waw9p5DU>

Conclusion

- We presented VIVO-Studio, a software assistant to the development of VIVO
- It is intended for professionals of multidisciplinary teams
- VIVO-Studio encapsulates the main tools necessary for the development of data processing functionality and ontology modeling
- VIVO-Studio has the possibility to be coupled to TopBraid Composer Free Edition, an ontology engineering tools.
<https://www.topquadrant.com/>
- About VIVO-Studio, future information will be updated on the UQAM VIVO-Wiki during the summer and indicated on the VIVO slack channel.
<https://wiki.uqam.ca/display/VIVOPUB>

View publication stats

Thank you!

Michel Héon

heon.michel@uqam.ca