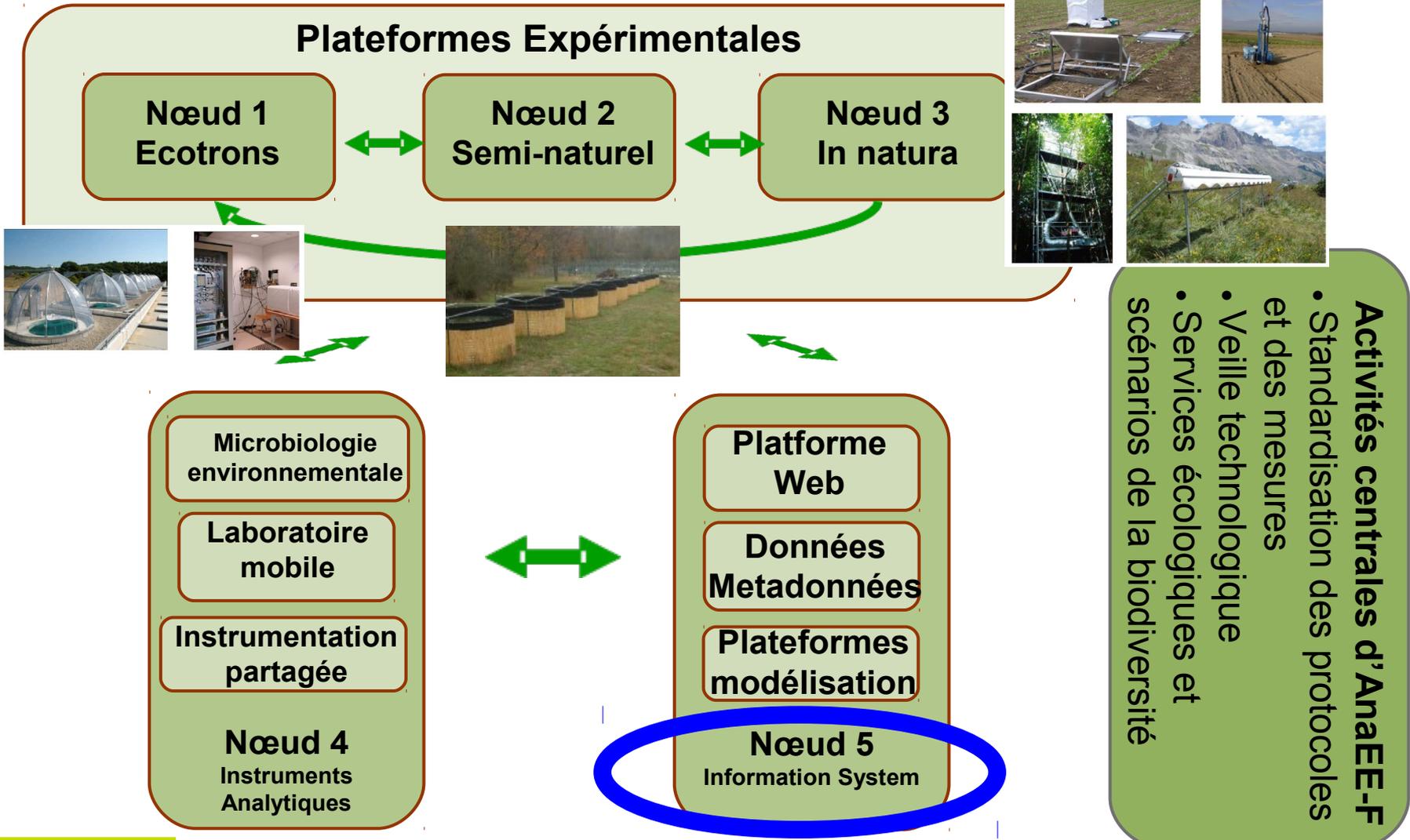


Interopérabilité au sein du SI AnaEE-France

AG du CATI SIOEA
19-20 mai 2015

C. Pichot, R. Yahiaoui et collectif



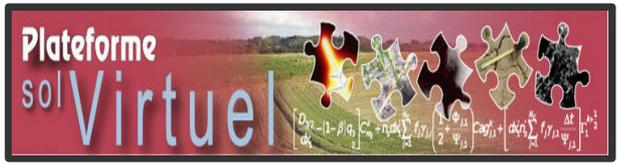
SI des SOERE (INRA)



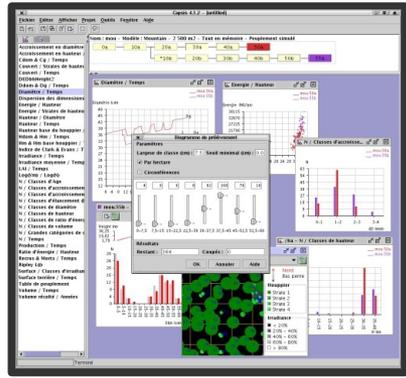
SI des experimentations à façon (CNRS-MNHN)




Plate-forme de modélisation et de simulation des agro-écosystèmes



Plateforme solVirtuel



Bases de données observ./expe.

Plateformes de modélisation

Une infrastructure distribuée pour l'observation/expérimentation :

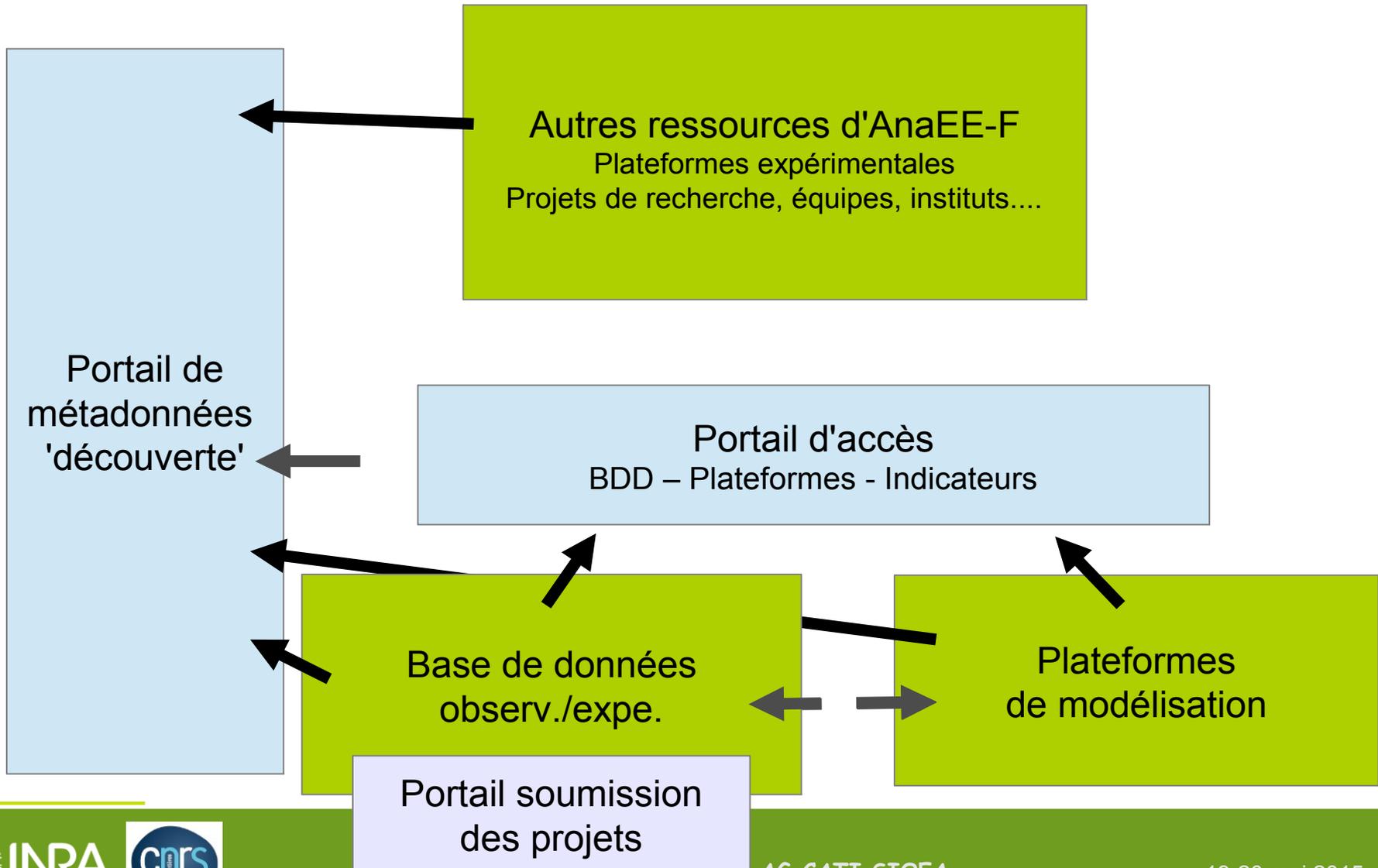
- sur des écosystèmes divers
- à plusieurs échelles
- comprenant la gestion des données
- offrant des services de modélisation

Un objectif de cohérence globale par

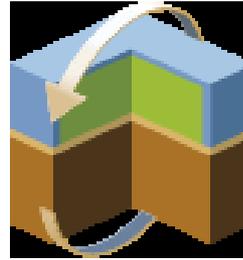
- la structuration, l'harmonisation
- le développement d'environnements de mutualisation
- un système d'information basé sur l'interopérabilité

Nécessitant notamment :

- définition et partage d'un vocabulaire
- Identification et qualification des ressources (via métadonnées)
- développement des services d'interopérabilité



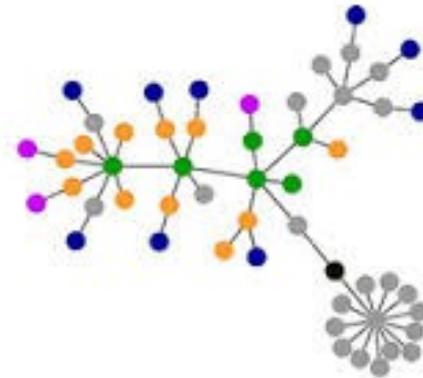
- Directe par webservice



SOERE ACBB



- Généralisée via la sémantique



Principe :

- Partager une **ontologie** de domaine modélisant les 'objets' et leur relation
- **Annoter** toutes les ressources (BDD, plateformes) sur cette ontologie
- Générer des unités élémentaires d'information sous forme de **triplets**
- **Requeter** en utilisant le référentiel sémantique

Mise en oeuvre dans d'AnaEE-France

* Approche basée sur des études de cas
(SI du SOERE OLA, Sol Virtuel, portail)

* S'appuie sur l'ontologie OBOE

* Mise en place d'outils partagés (gestion de l'ontologie et annotation)

ECOLOGICAL INFORMATICS 2 (2007) 279-296

available at www.sciencedirect.com

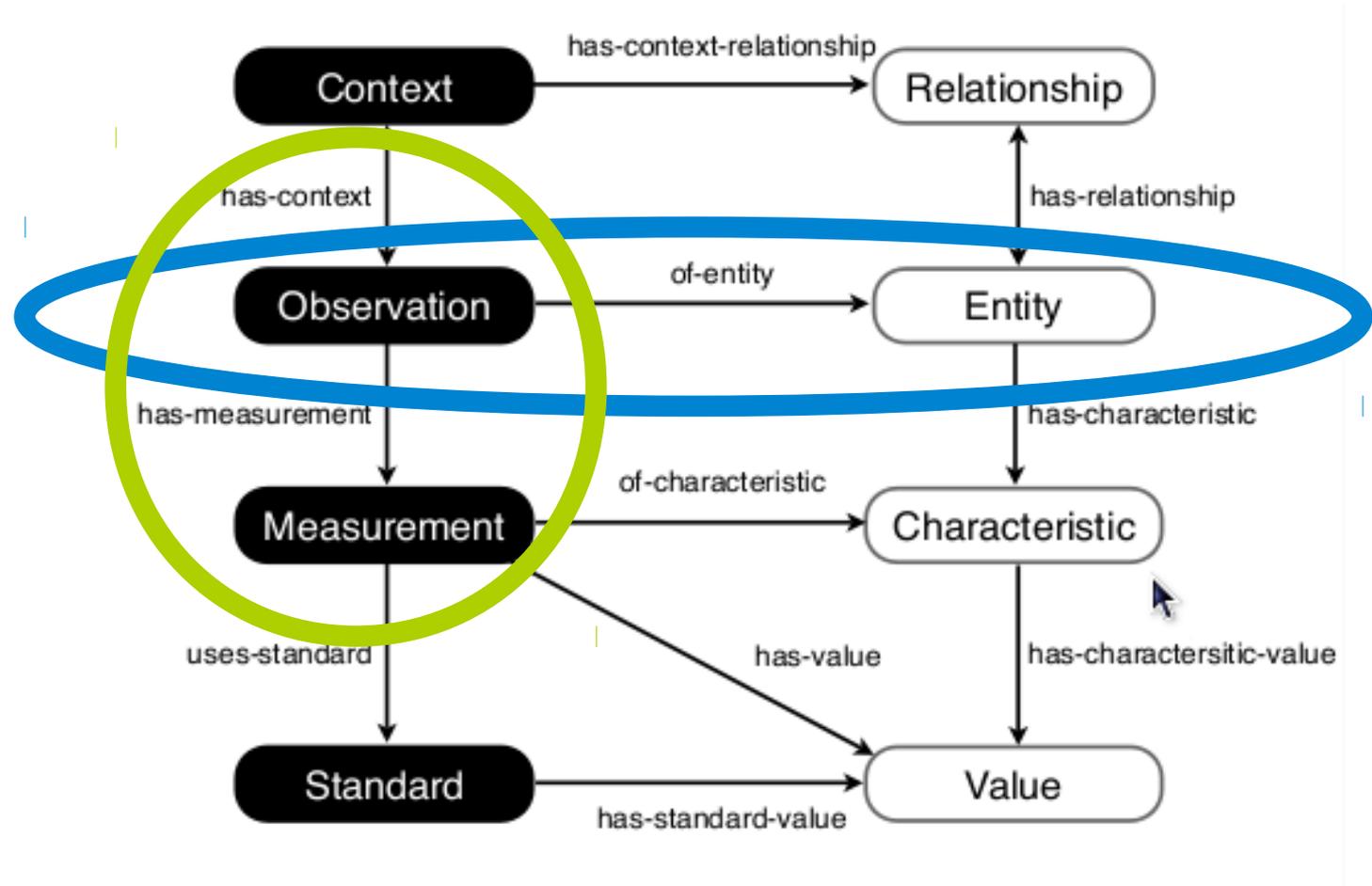
ScienceDirect
www.elsevier.com/locate/ecolinf

An ontology for describing and synthesizing ecological observation data

Joshua Madin^{a,e,*}, Shawn Bowers^b, Mark Schildhauer^a, Serguei Krivov^c,
Deana Pennington^d, Ferdinando Villa^c

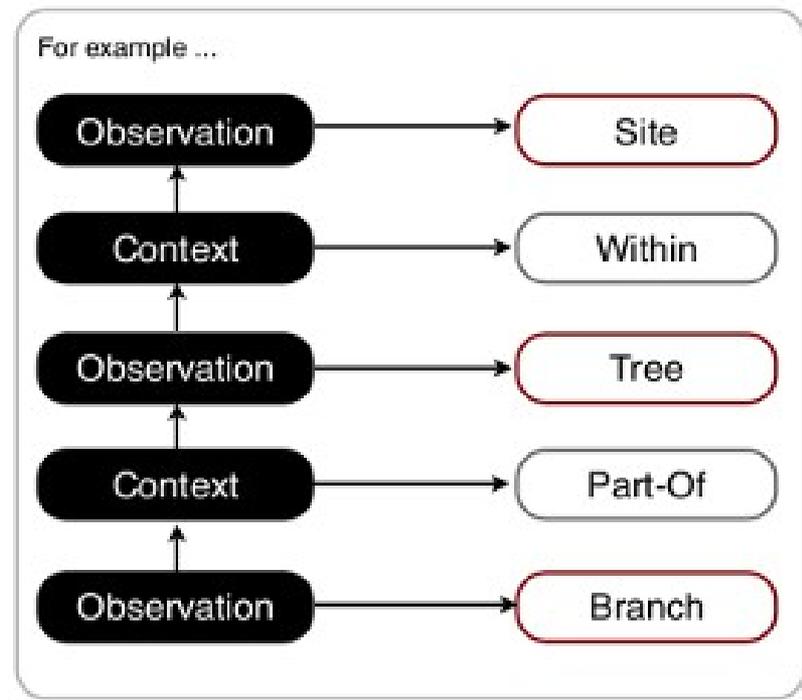
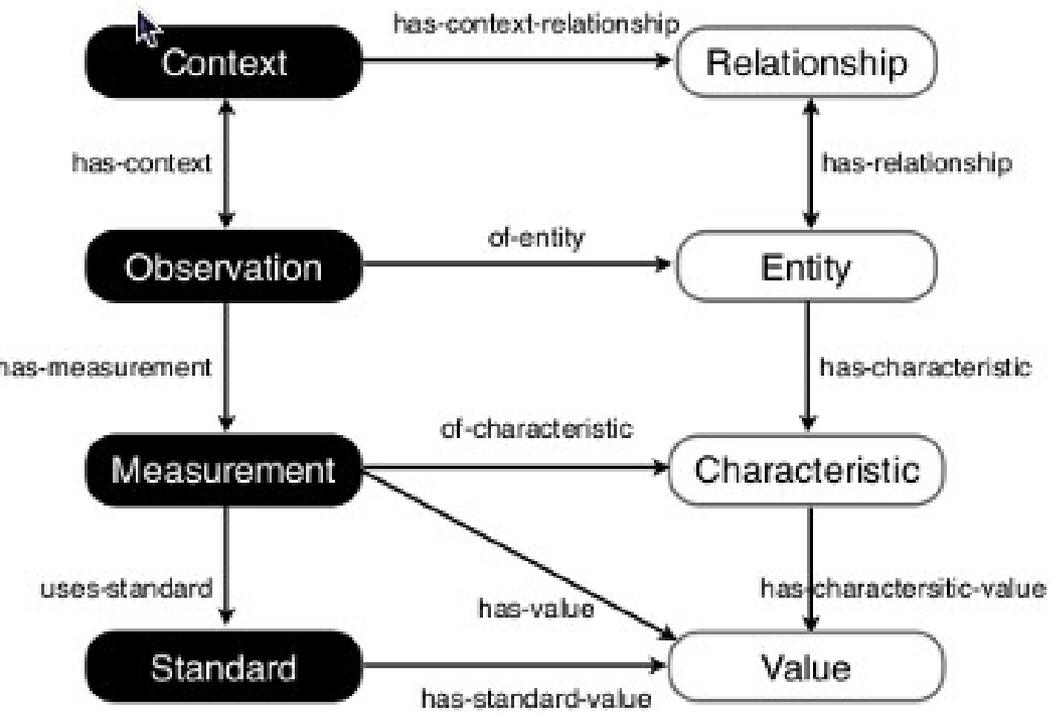
^aNational Center for Ecological Analysis and Synthesis, University of California, Santa Barbara, California 93101, USA
^bGenome Center, University of California, Davis, California 95616, USA
^cGund Institute for Ecological Economics, University of Vermont, Burlington, Vermont 05405, USA
^dUniversity of New Mexico, Albuquerque, New Mexico 87131, USA
^eDepartment of Biological Sciences, Macquarie University, New South Wales 2109, Australia

L'ontologie OBOE



The ontology separates observations from the entity being observed: the observation has a measurement while the entity has characteristics, and the measurement is then of that characteristic.

L' « observation » dans OBOE



Observations can occur within a context, which in turn is an observation;
this property is transitive.



Interopérabilité

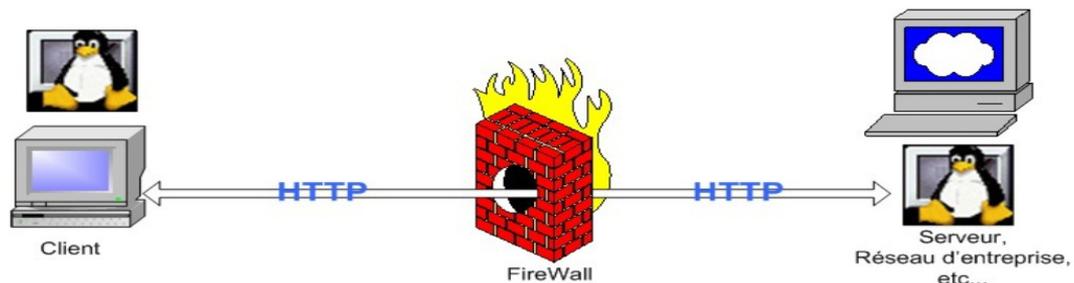


SOERE <-> Plate-forme de modélisation
ACBB - Plates-formes Sol-Virtuel

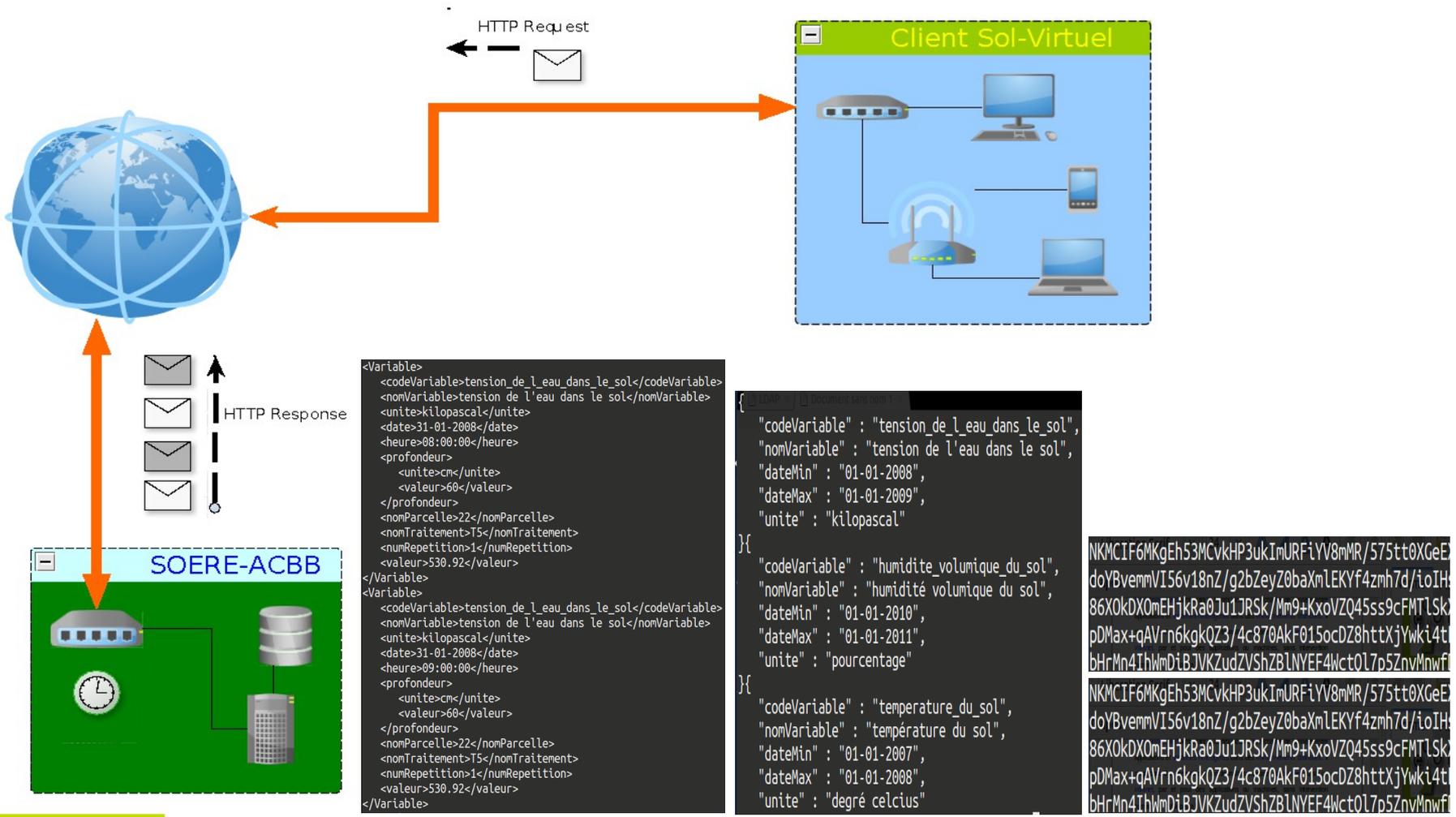
ANAEE-France
Web-Sémantique (Ontologie)

Web-Services

Un service web est un **programme informatique** de la famille des **technologies web** permettant la communication et l'échange de données entre applications et **systèmes hétérogènes** dans des **environnements distribués**. Il s'agit donc d'un ensemble de fonctionnalités exposées sur **internet** ou sur un **intranet**, par et pour des applications ou machines, sans intervention humaine... (Wikipedia)



SOERE-ACBB ↔ Plate-forme Sol-Virtuel



Choix de l'architecture web service

SOAP (SOA)	REST (ROA)
<p>Standard (Protocole générique pour le transport : http, smtp, jms).</p> <p>Verbeux (XML, SEI)</p> <p>Contrats formels</p> <p>Opérations avec ou sans état (stateless)</p>	<p>Opérations sans état (stateless)</p> <p>Protocole HTTP pour le transport</p> <p>Accessible, Facile à mettre en œuvre</p> <p>Utilise les verbes standards du protocole HTTP pour faire du CRUD (GET PUT POST DELETE)</p>

Critères de Qualité



Flexibilité



Robustesse



Performance



Réutilisabilité

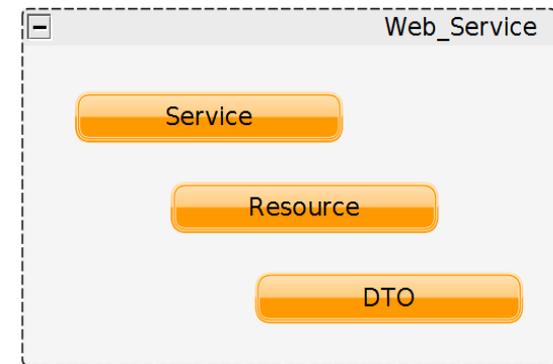
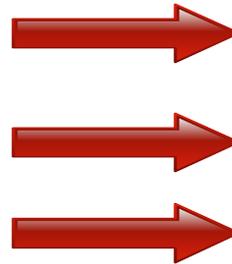


Sécurité



Flexibilité

Un web service est une « *Fonction* » capable d' **intercepter** des requêtes HTTP, d'en extraire la demande, de l'**exécuter**, puis **formater** le résultat pour enfin le renvoyer au client.



Package service web



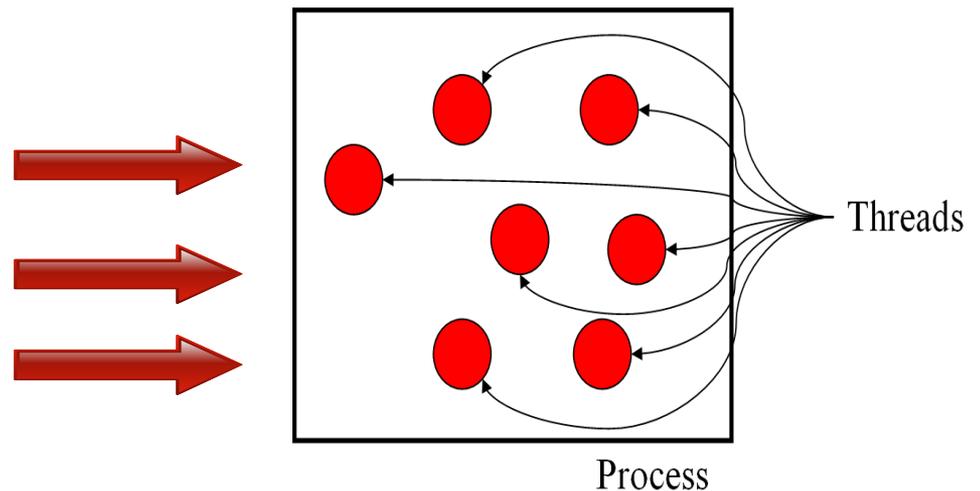
Performances (Temps de réponse)

Une thread (appelée aussi processus léger) est un fil d'instructions (un chemin d'exécution) à l'intérieur d'un processus.

Contrairement aux processus, les threads d'un même processus partagent le même espace d'adressage et le même environnement,

~~Dénormalisation DB~~

VS



Débit d'extraction SANS // ~ 0,5 Mo/s

Débit d'extraction AVEC // ~ 9 Mo/s



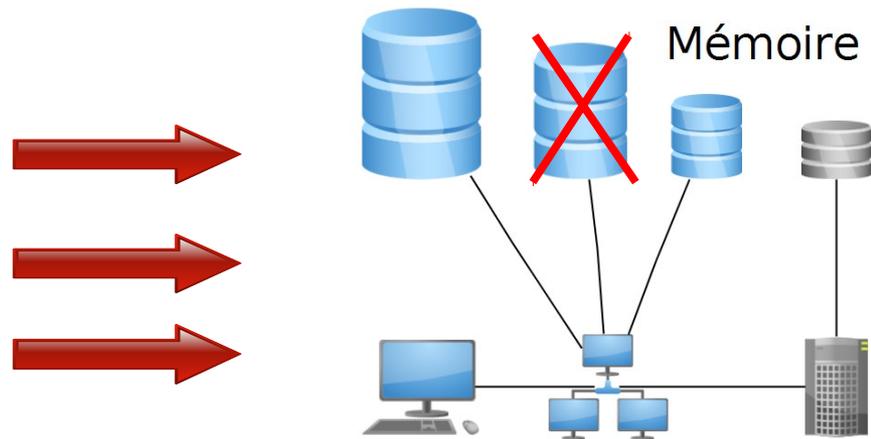
Robustesse

Contrôler la charge mémoire
(empreinte mémoire) afin
d'éviter toute défaillance du
système.

Pagination des requêtes SQL

LIMIT – OFFSET

Volume de données >> à la mémoire

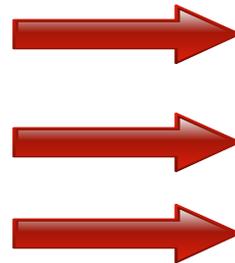




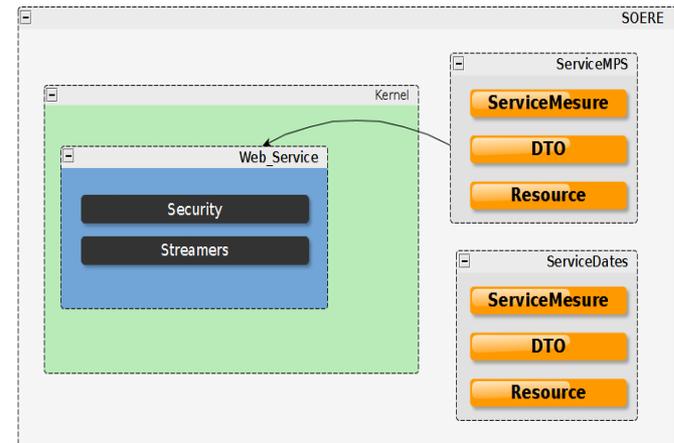
Réutilisabilité

Intégration du web service
(core) dans le Kernel SOERE.

Implémentation des web
services (spécifiques à chaque
SOERE) dans la couche plugin
(basée directement sur le web
service core)



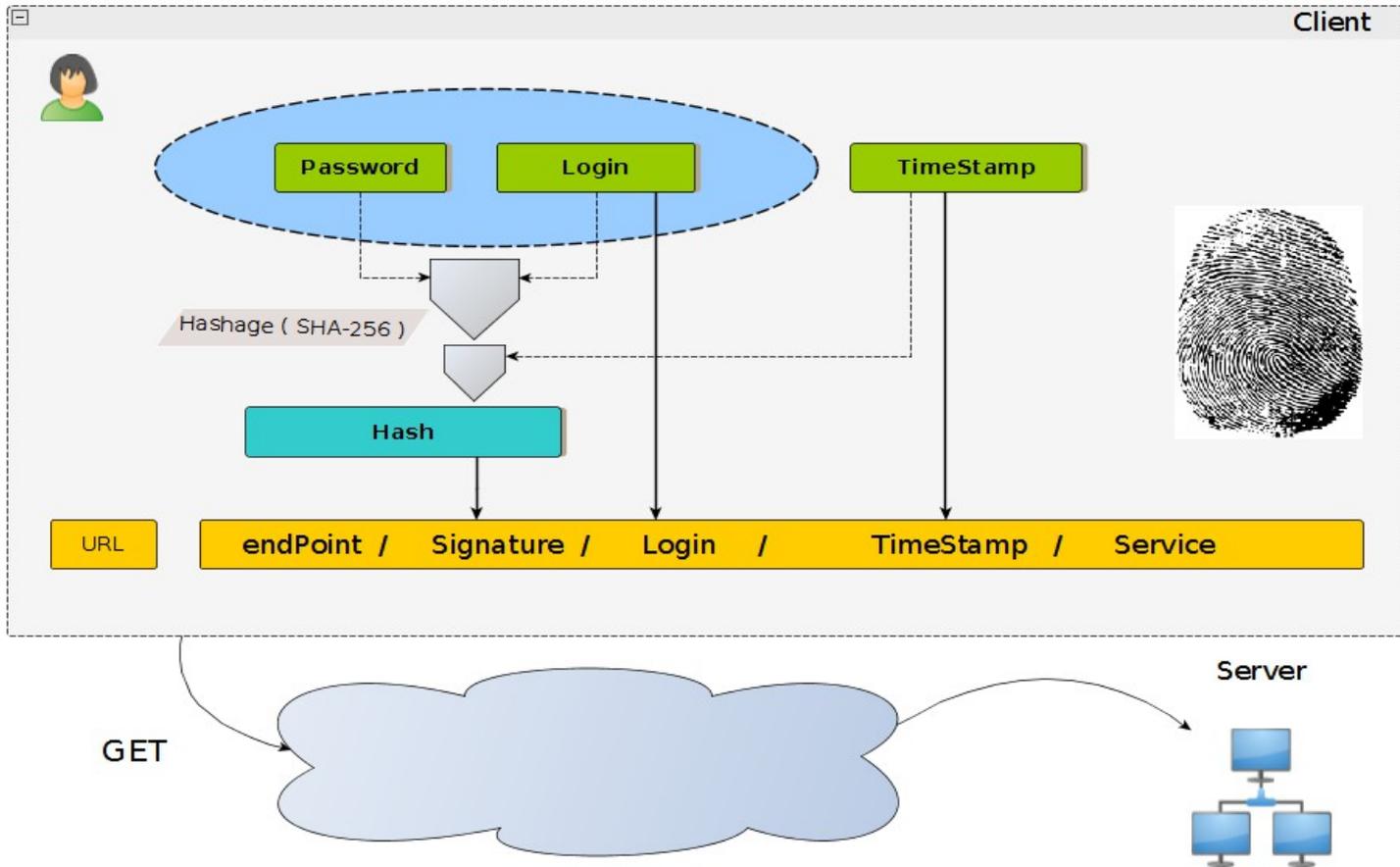
Write once, run anywhere





Sécurité (1/2) ~~SSL~~

Client

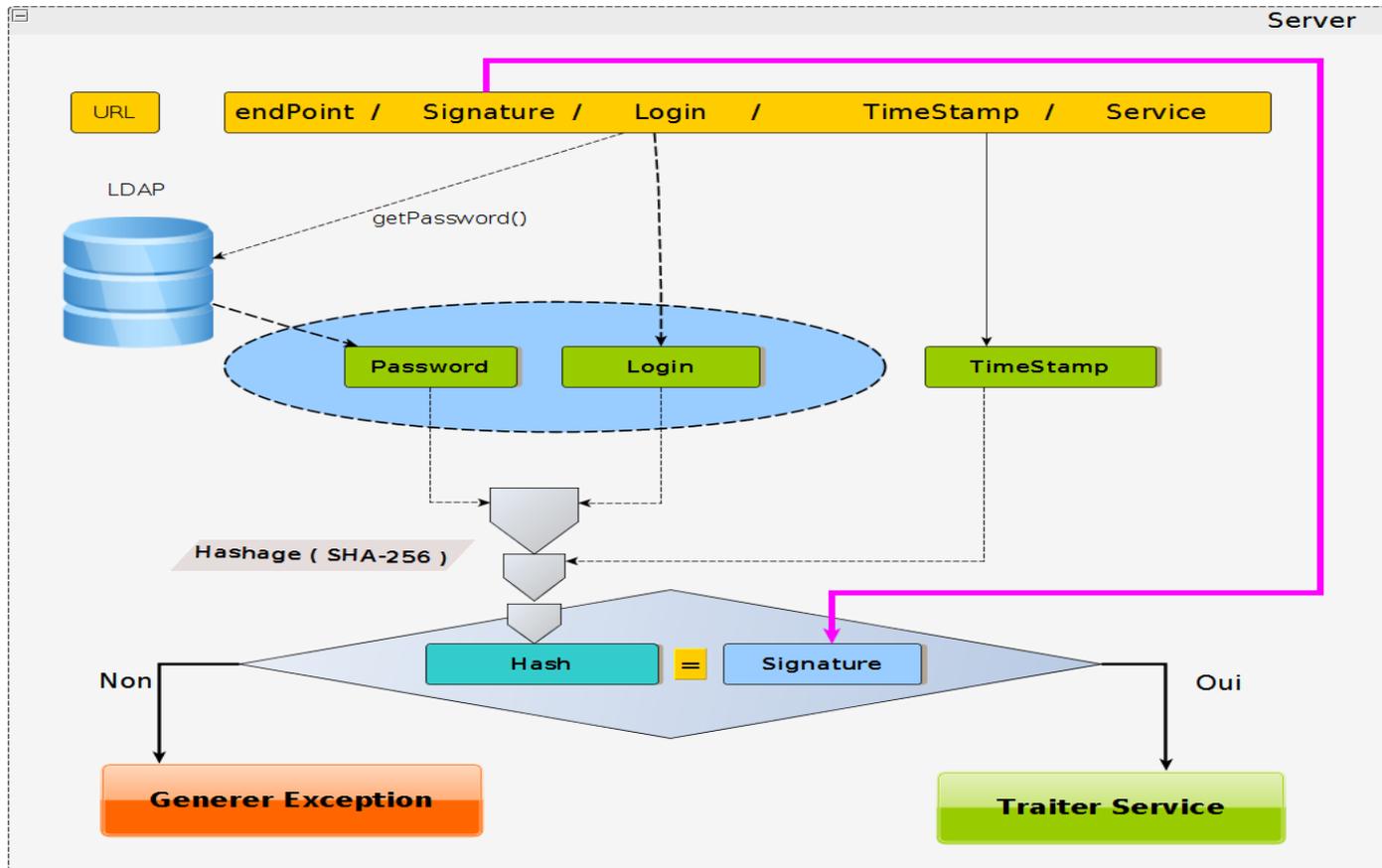




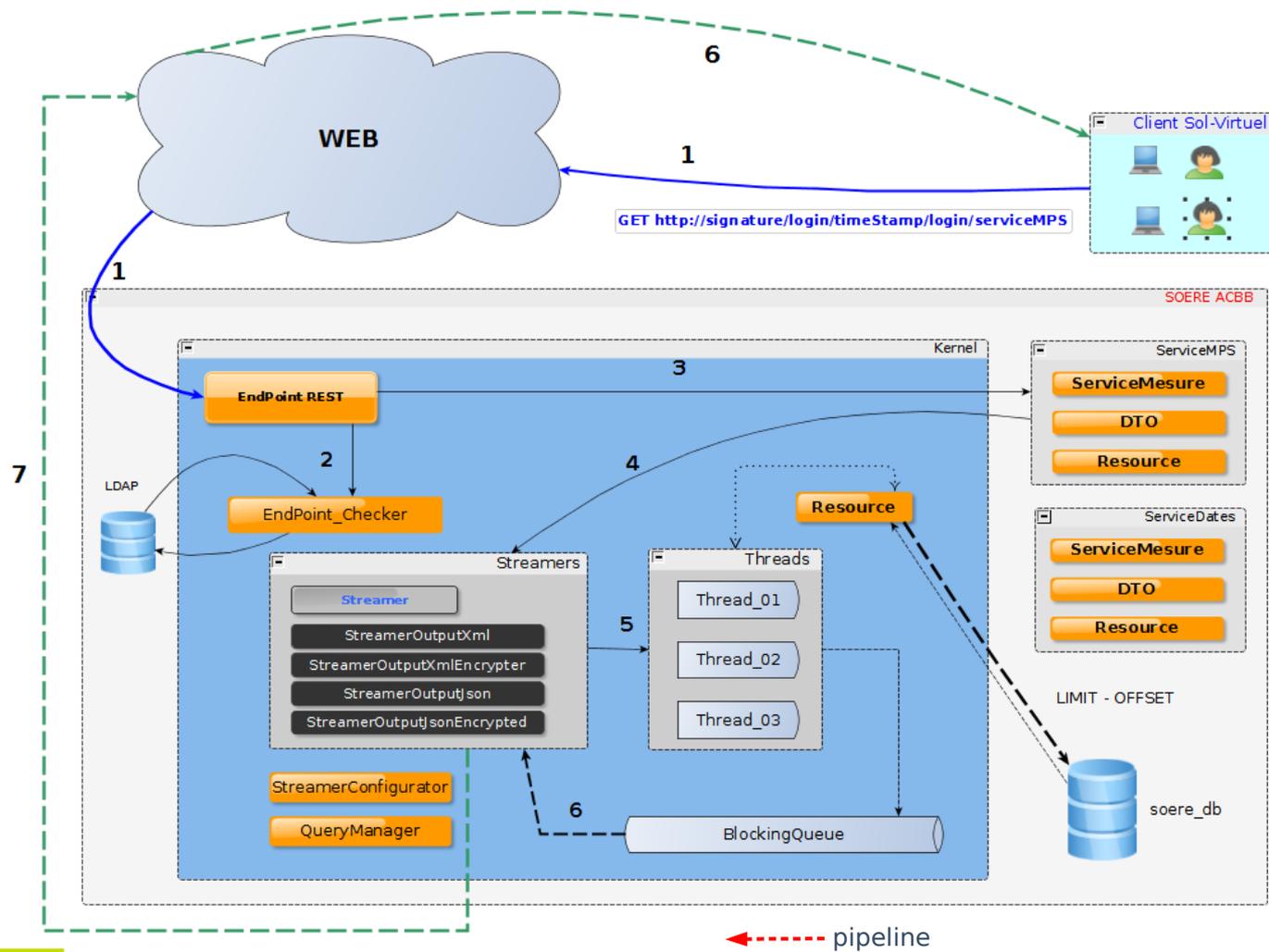
Sécurité (2/2)

~~SSL~~

Serveur

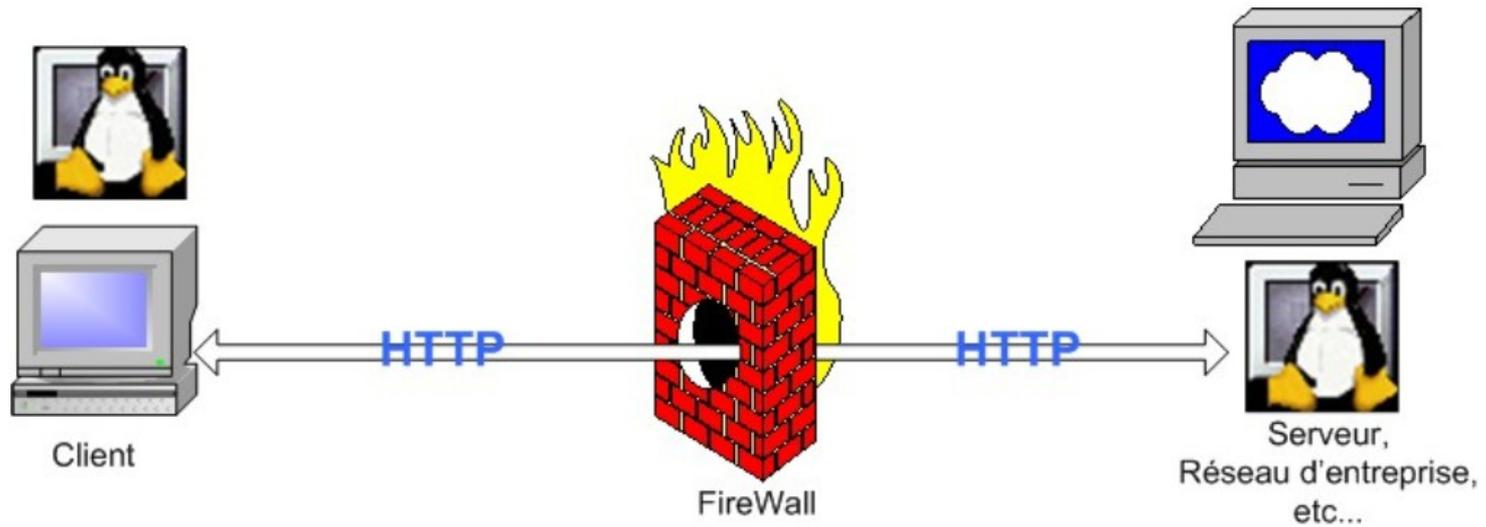


Fonctionnement du système



← pipeline

La Théorie



La Réalité

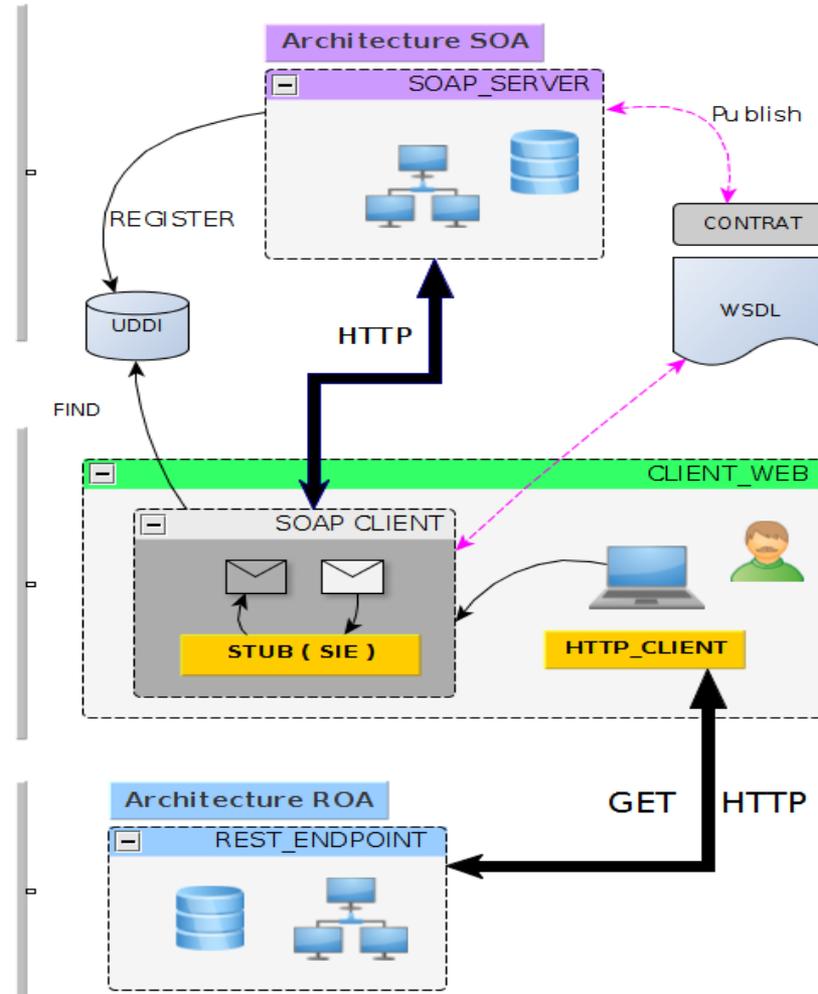
DATA_SOAP_SERVER		
TMP	DT	M_TPM

TEMP = TMP ??

DATA_CLIENT		
TEMP	TJ	TMP_M

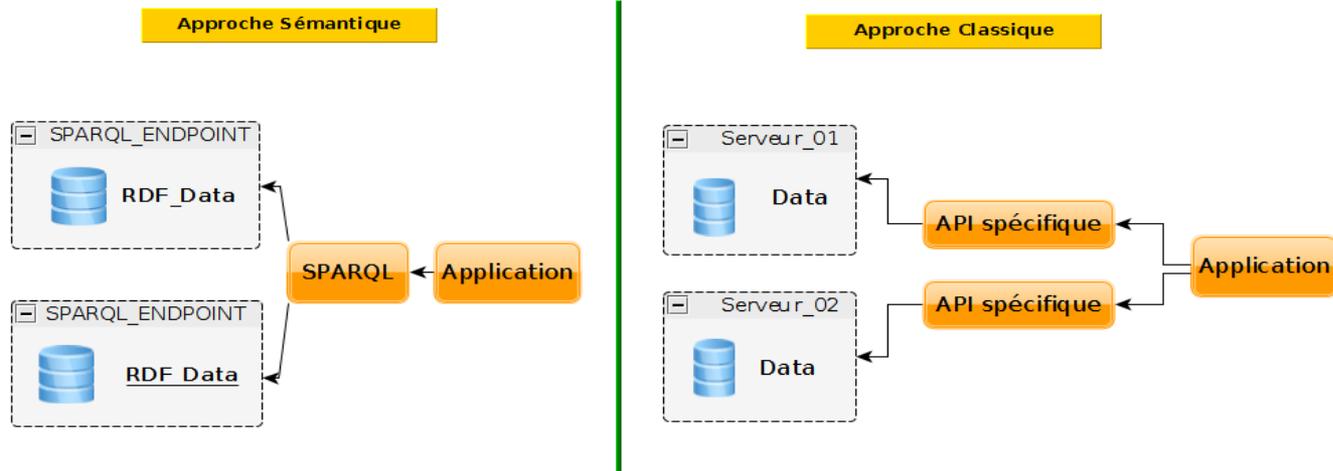
daily_TMP = TJ ??

DATA_REST_ENDPOINT	
monthly_tmp	daily_TMP

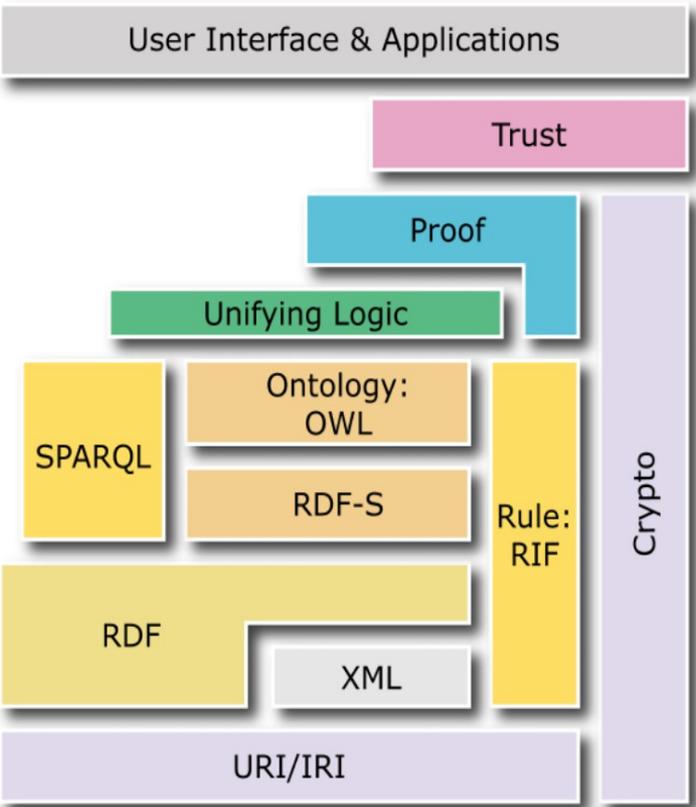




Le Web sémantique, ou toile sémantique, est un mouvement collaboratif mené par le World Wide Web Consortium (W3C) qui favorise des méthodes communes pour échanger des données sur Internet pour accéder simplement.. (Wikipedia)

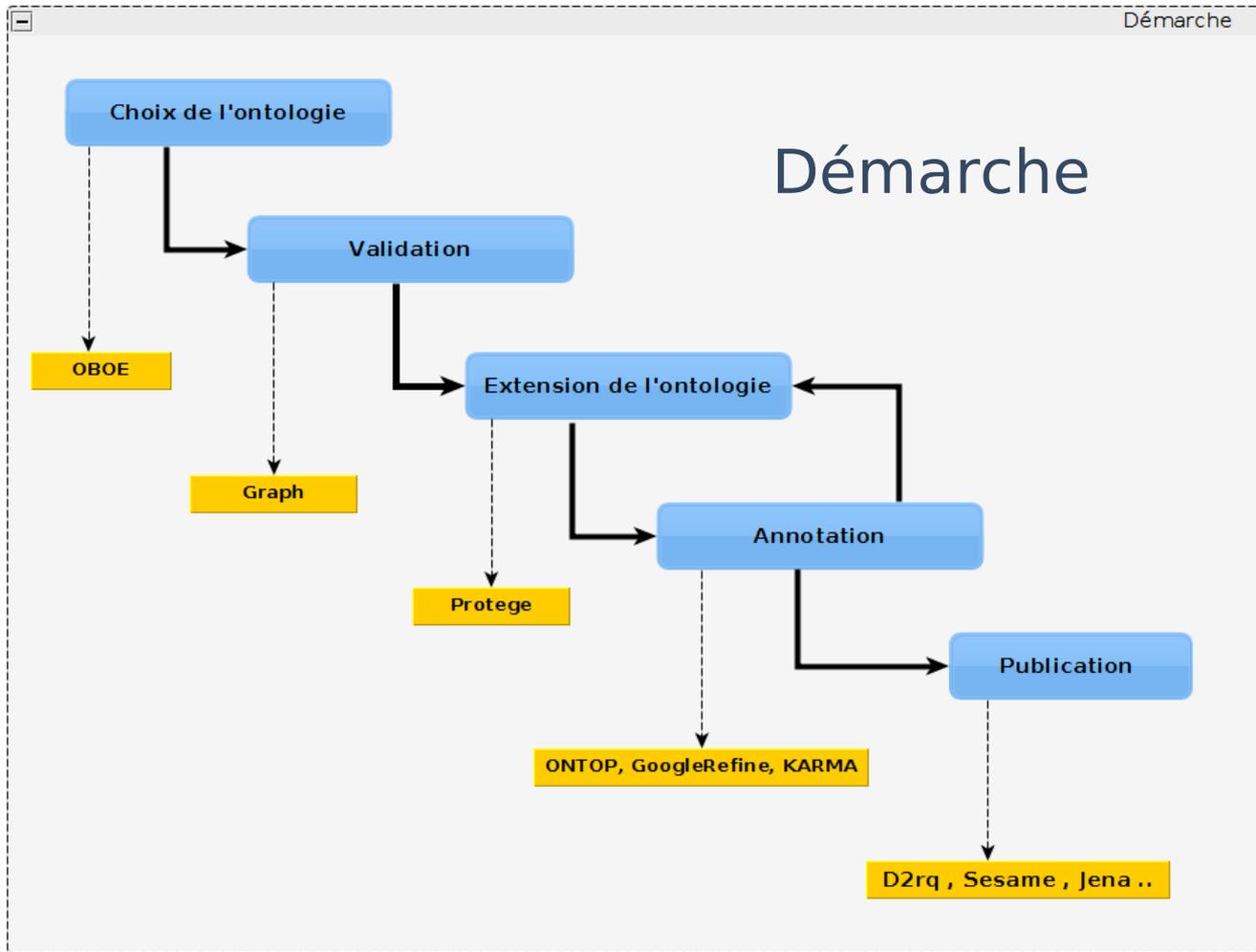


Pile de Standardisation



Pile des standards du Web de données W3C®

Mise en place de l'ontologie - AnaEE-F



Choix de l'ontologie

Oboe ??

Ontologie conçue
comme étant un
modèle générique
pour la
modélisation et la
représentation des
observations
scientifiques

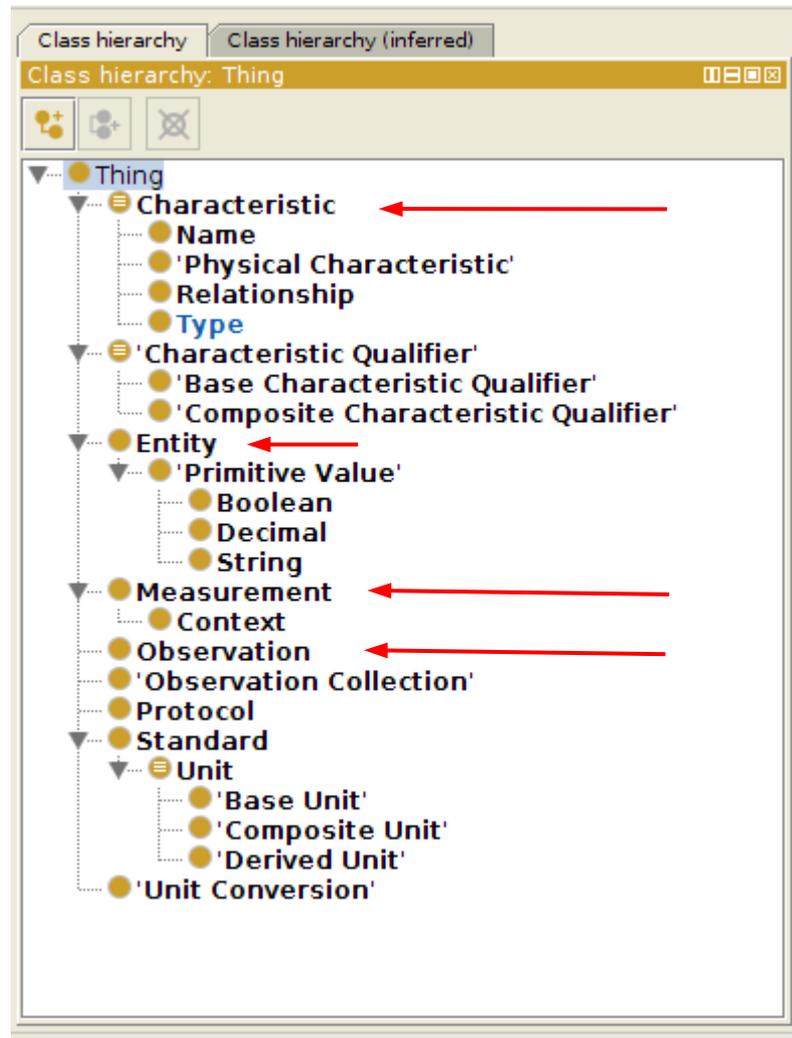
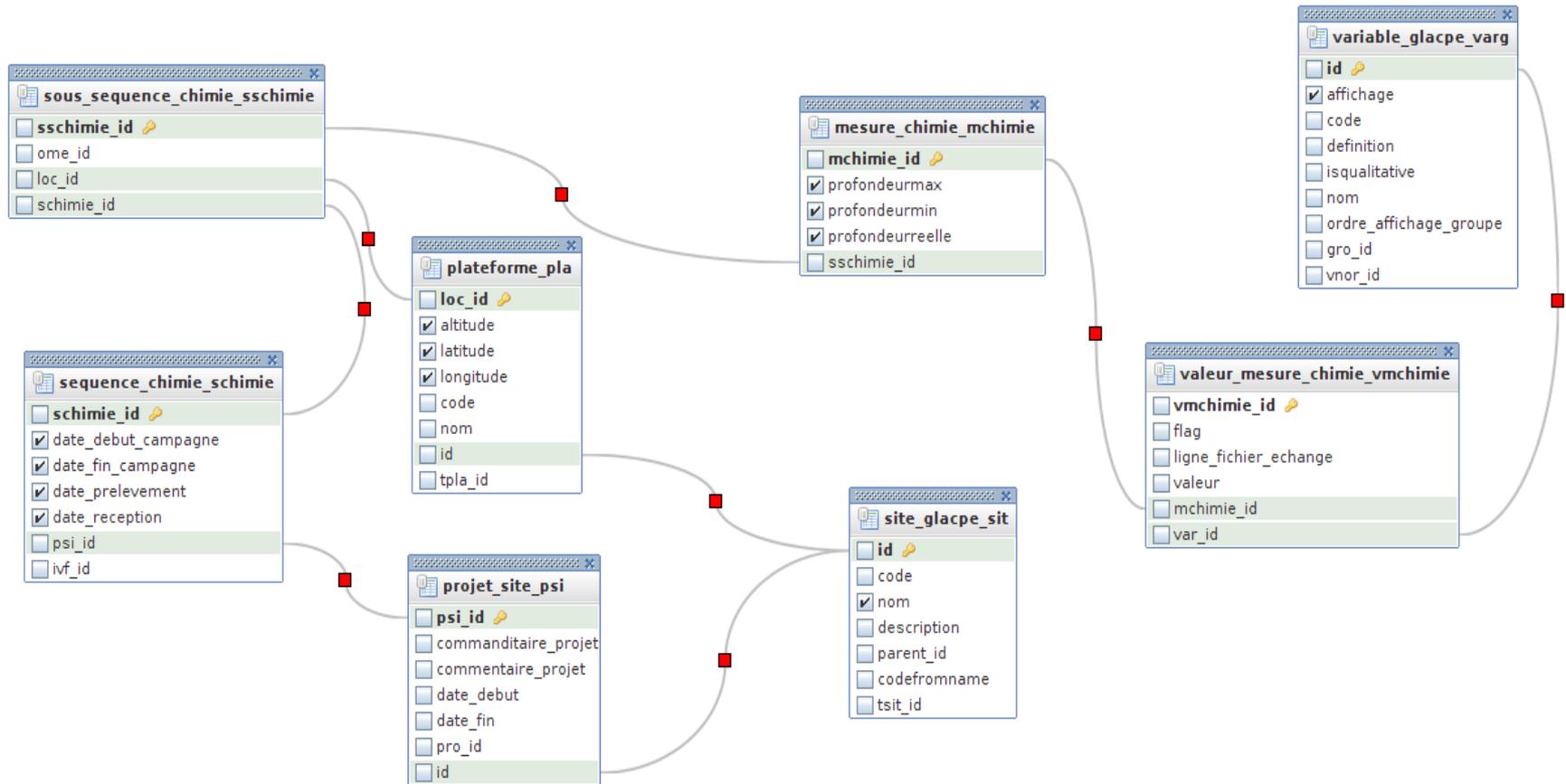


Schéma cas Physico-Chimie de OLA

Validation



Données OLA

Validation

Localisation des entités

	A	D	E	F	G	H	I	J	K	L
1	Banks Isl									
2		MG/L	MG/L	MG/L	MG/L	MG/L	MG/L	MG/L	MG/L	MG/L
3		Na	K	SO4	Cl	Al	Ba	Fe	Li	M
28	BKY	1,7	1,0	9,30	2,64	0,990	0,0341	1,130	0,0030	0,020
29	BKZ	232,0	10,8	68,70	404,00	0,110	0,0303	0,644	0,0140	0,145
30	BKAA	3,3	0,3	2,20	12,20	0,020	0,0181	0,230	0,0005	0,020
31	BKAB	2,7	0,7	0,10	8,46	0,010	0,0292	0,068	0,0010	0,000
32	BKAC	22,1	3,1	13,40	72,80	1,230	0,1250	4,940	0,0090	0,220
33	BKAD	1,9	0,7	34,40	3,91	0,220	0,0159	0,508	0,0020	0,010
34	BKAE	3,5	0,5	32,90	17,10	0,010	0,0139	0,049	0,0005	0,000
35	BKAF	0,9	0,3	5,00	2,52	0,005	0,0126	0,034	0,0010	0,000
36	BKAG	1,1	0,2	9,40	2,35	0,010	0,0103	0,075	0,0005	0,000
37	BKAH	0,4	0,3	1,60	0,84	0,010	0,0072	0,054	0,0005	0,020
38	BKAI	0,5	0,4	4,70	1,22	0,170	0,0061	0,304	0,0010	0,020
39	BKAJ	1,1	0,1	26,00	1,50	0,030	0,0206	0,080	0,0010	0,000

Site

Observation

Oui mais.. pas suffisant !

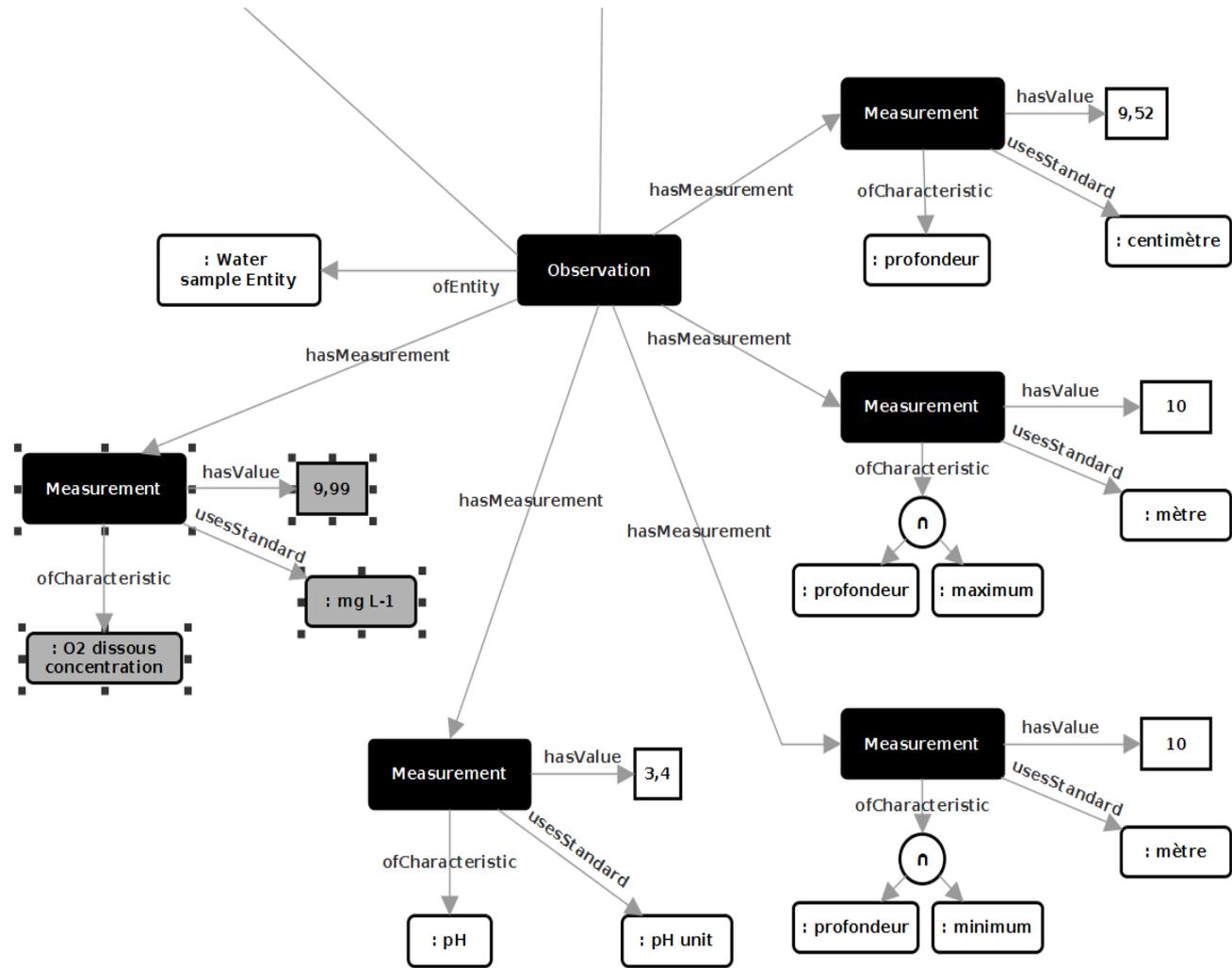
	AA	AB	AC	AD	AE	AF	AG
		uS	Celsius	m asl			L=Lake
	PH	COND	TEMP	ELEV	LAT	LONG	P=Pond
50	7,6	48,0	6,0	2	74 30.69N	121 41.08W	L
50	8,1	1160,0	8,0	0	74 27.82N	122 34.55W	P
50	7,6	83,0	5,0	8	74 21.46N	124 33.92W	P
50	8,1	89,0	7,0	20	74 08.10N	124 12.52W	P
50	8,4	333,0	8,0	0	72 21.13N	125 24.43W	P
50	7,8	137,0	3,0	122	71 43.79N	123 28.94W	L
50	8,4	216,0	7,5	169	71 34.48N	122 30.69W	P
50	7,8	109,0	3,5	175	72 14.22N	121 29.96W	L
50	7,9	105,0	8,0	105	72 39.96N	119 56.11W	P
50	7,7	41,0	3,0	131	73 35.57N	119 35.01W	L
50	7,9	65,0	4,0	137	73 20.79N	116 46.23W	L
50	8,5	137,0	7,0	195	73 29.11N	115 41.15W	P

PH

Measurement

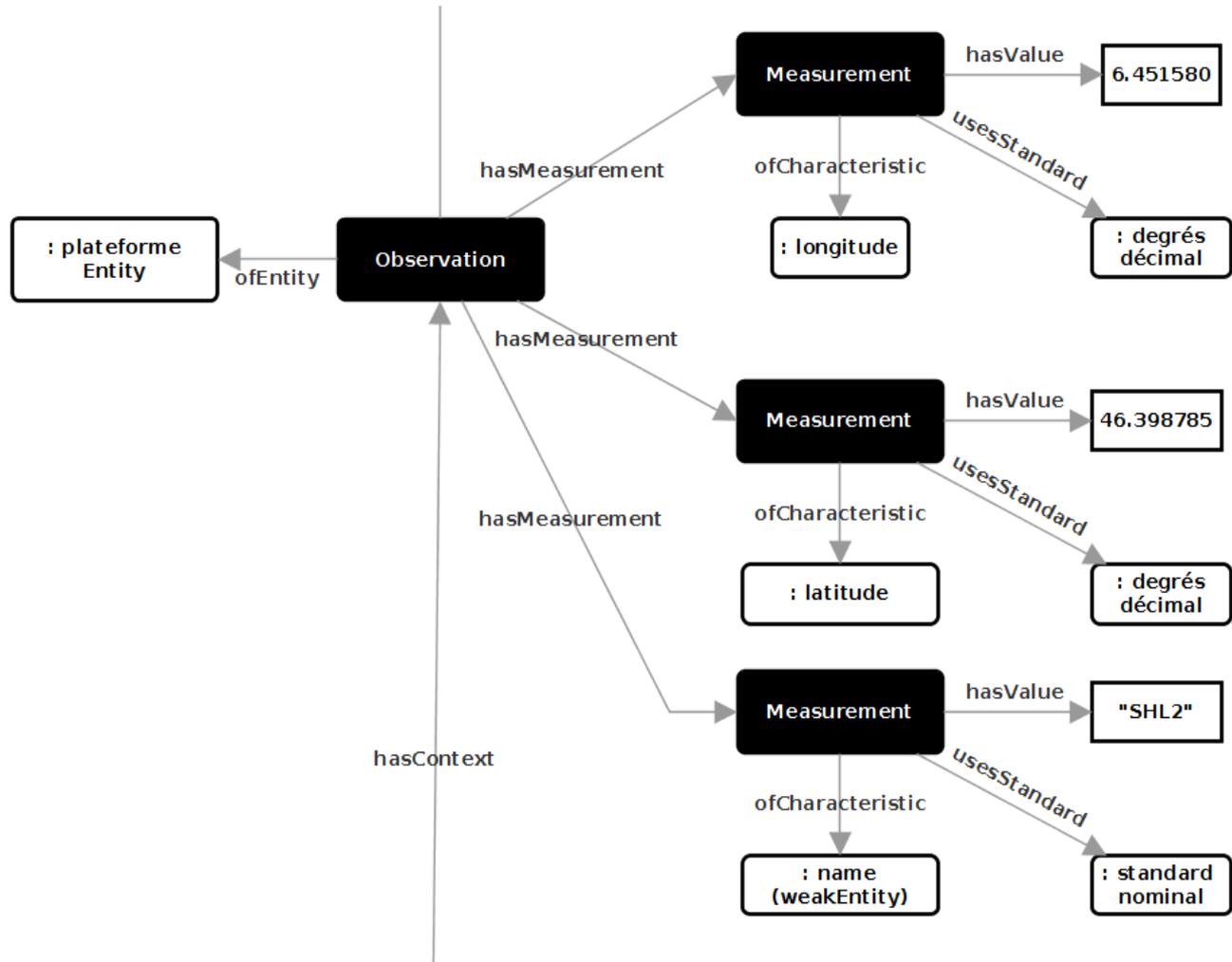
Location

Graph RDF 1/3



Validation

Graph RDF 2/3

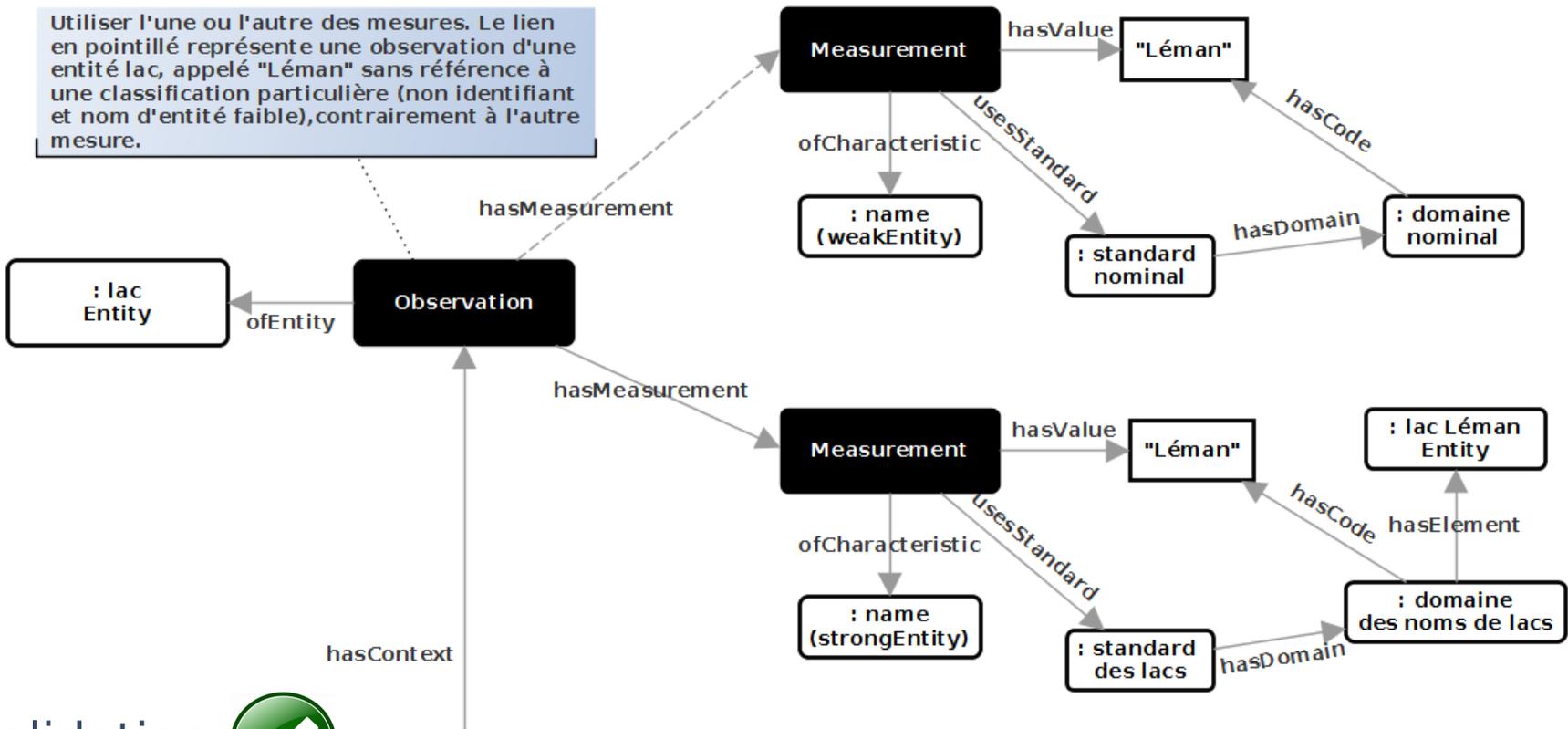


Validation

Graph RDF

3/3

Utiliser l'une ou l'autre des mesures. Le lien en pointillé représente une observation d'une entité lac, appelé "Léman" sans référence à une classification particulière (non identifiant et nom d'entité faible), contrairement à l'autre mesure.

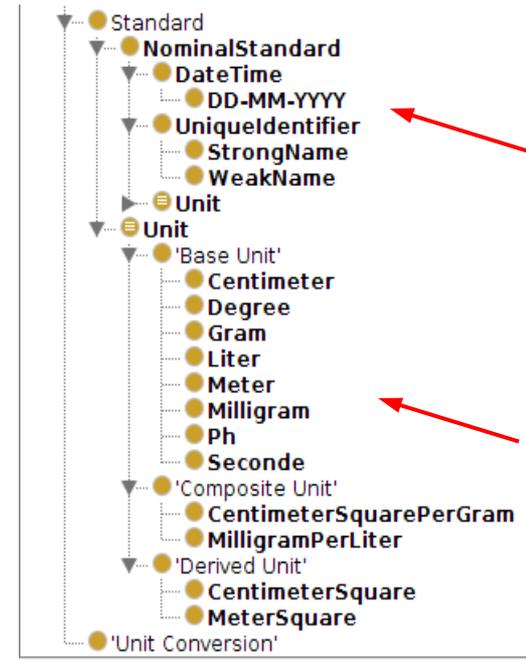
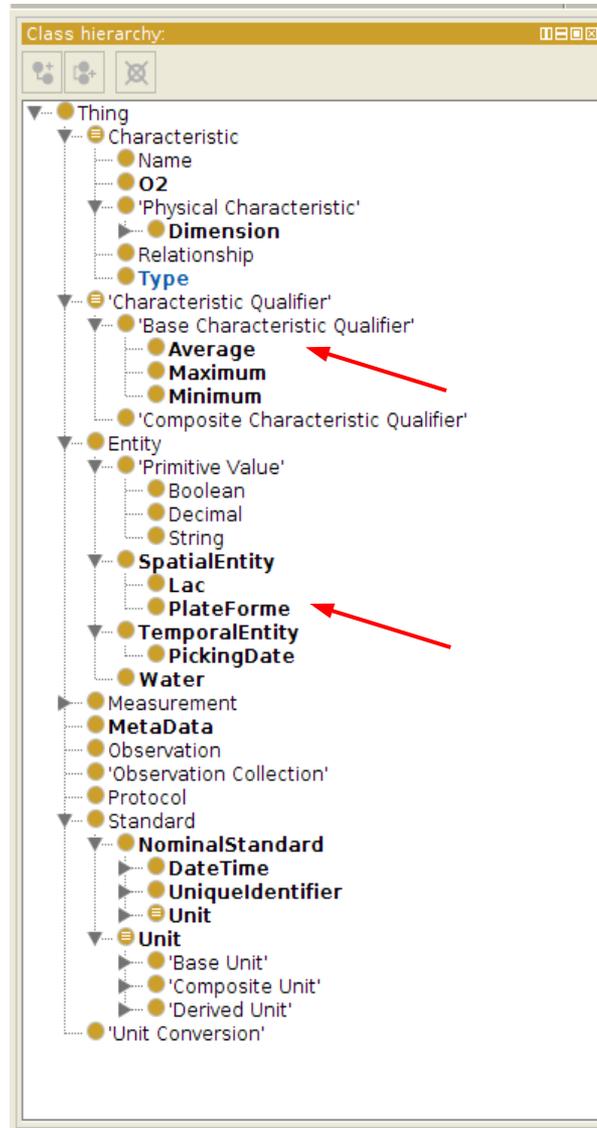


Validation



Extension OBOE

OBOE-CORE
+
'Thésaurus'



Transformation RDB – RDF

Prèlevement eau		
code_échantillon	PH	profondeur
P10	6.5	12.5



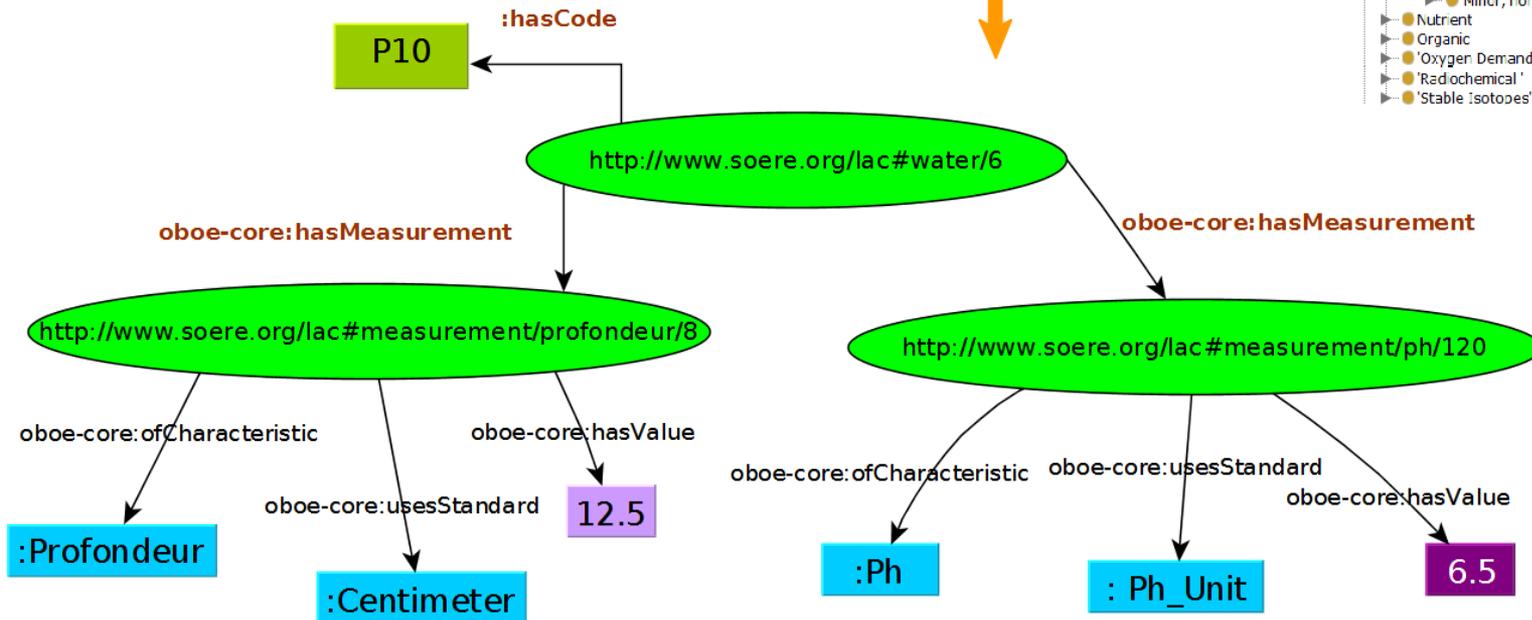
R2RML



correspondances

Schéma

- Chemical
 - Inorganic
 - 'Dissolved Gas'
 - 'Dissolved Solids'
 - Major
 - Minor
 - 'Mincr, metals'
 - 'Mincr, non-metals'
 - Nutrient
 - Organic
 - 'Oxygen Demand'
 - 'Radiochemical'
 - 'Stable Isotopes'



Outils

D2RQ

Web
Karma



ontop



Google Refine

Stratégies de choix du Stockage

Stockage des triplets RDF dans un tripleStore :

- Allège la charge pesante sur la BD-R
- Duplication des données !
- Triplets RDF **désynchronisés** avec la BD-R

Mapping à la volée :

- Charge supplémentaire pour la BD-R
- Triplets **Synchronisés** avec la BD-R

Schéma BD OLA. (Cas Physico-Chimie)

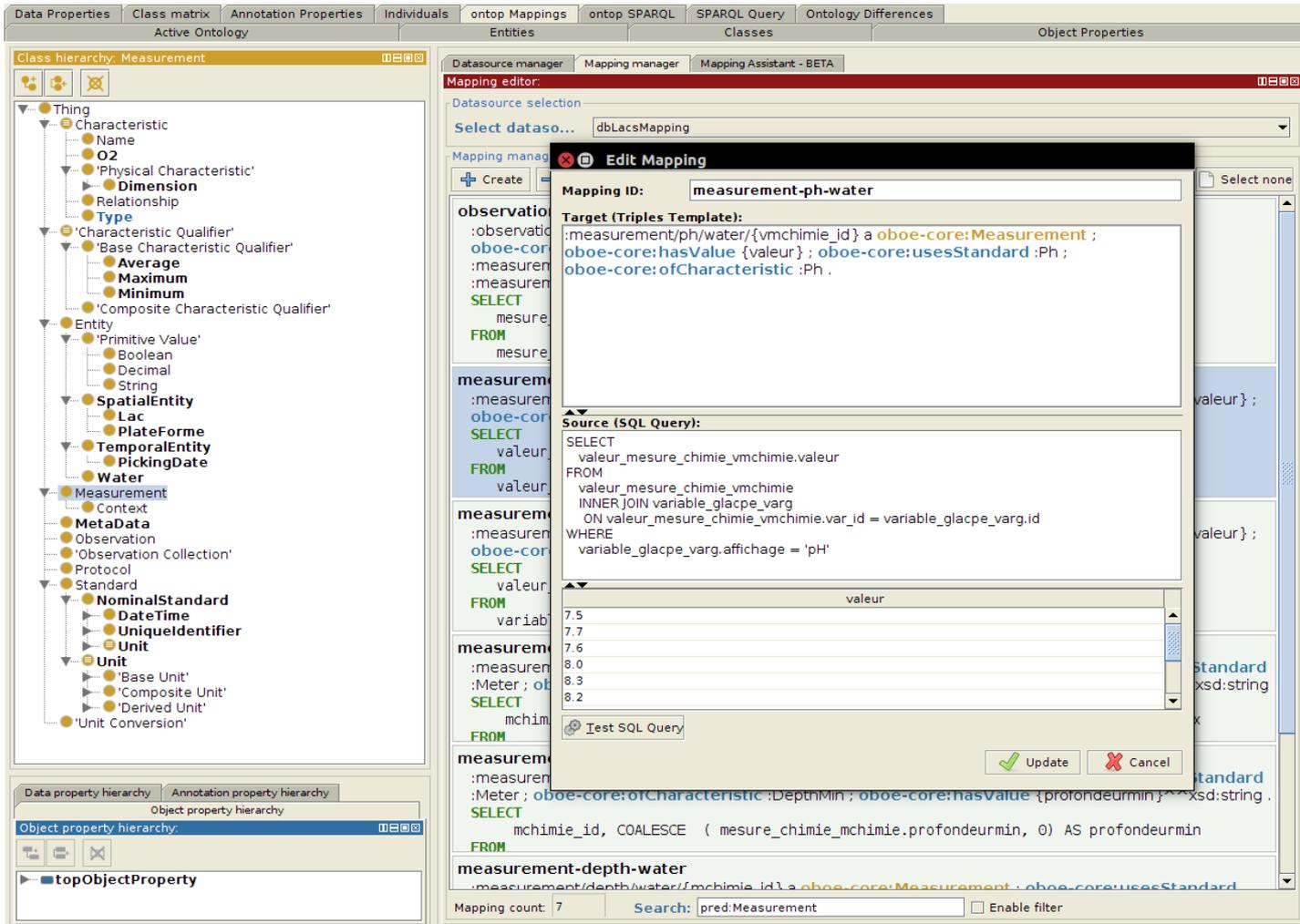
Attributs	Water				PlateForme			Lac		Date_Prélèvement
Entity	:Water				:PlateForme			:Lac		:TemporalPoint
Caractéristiques	Profondeur	Prof Min	ph	O2	Name	:Spacial_Localisation Latitude Longitude		weakName	StrongName	Date
Measurements Standard	Centimètre	Mètre	ph	Mg L-1	Nominal	Degrès	Degrès	Nominal	Nominal	Date_Time
Précision	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Colonne	Profondeur	Prof_Min	PH	O2	P.F.Name	Latitde	Longitude	WeakName	StrongName	Date_Prélèvement
Valeur	9.53	10	3.4	9.99	SHL2	46.39	6.45	Léman	Léman	28-05-2009

Mapping

Le mapping se fait en deux temps

- 1- Définition d'un Triples Template (syntaxe Turtle)
- 2- Définition du Source (SQL Query)

Mapping des mesures : (PH)



The screenshot shows the Protege software interface with the 'Mapping editor' window open. The 'Class hierarchy' on the left shows the 'Measurement' class and its subclasses, including 'Dimension', 'Type', 'Entity', and 'Unit'. The 'Mapping editor' window displays the following information:

- Mapping ID:** measurement-ph-water
- Target (Triples Template):**

```

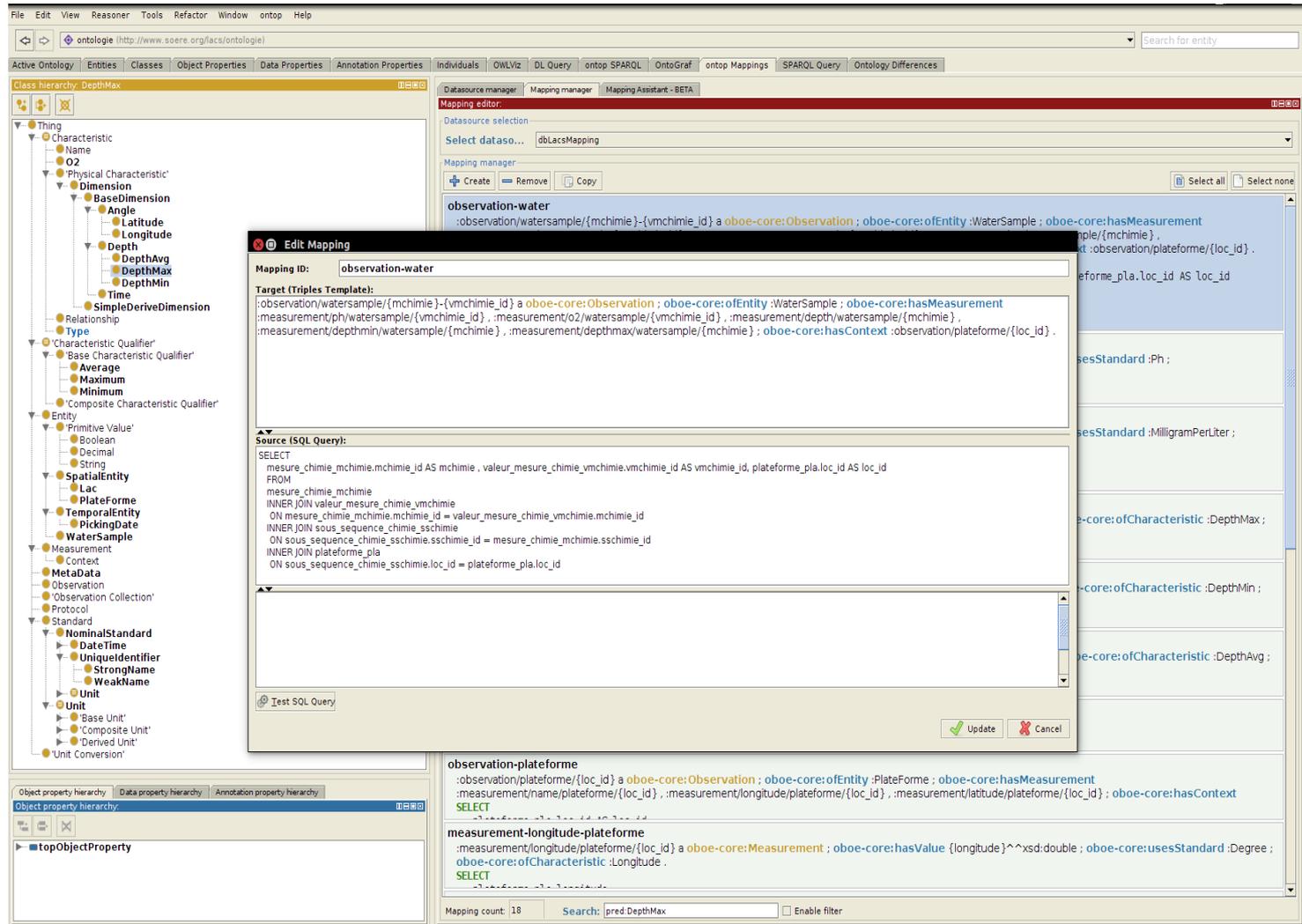
:measurement/ph/water/{vmchimie_id} a oboe-core:Measurement ;
oboe-core:hasValue {valeur} ; oboe-core:usesStandard :Ph ;
oboe-core:ofCharacteristic :Ph .
        
```
- Source (SQL Query):**

```

SELECT
  valeur_mesure_chimie_vmchimie.valeur
FROM
  valeur_mesure_chimie_vmchimie
  INNER JOIN variable_glacpe_varg
    ON valeur_mesure_chimie_vmchimie.var_id = variable_glacpe_varg.id
WHERE
  variable_glacpe_varg.affichage = 'pH'
        
```
- Table:** A table with one column 'valeur' containing the following values: 7.5, 7.7, 7.6, 8.0, 8.3, 8.2.

At the bottom of the mapping editor, there are 'Update' and 'Cancel' buttons, and a 'Test SQL Query' button. The 'Mapping count' is 7, and the search filter is 'pred:Measurement'.

Mapping des observations : (Water)



The screenshot shows the Protégé ontology editor interface. On the left, a class hierarchy tree is visible, with 'DepthMax' selected under 'Dimension'. The main workspace displays the 'Edit Mapping' dialog for the mapping ID 'observation-water'. The dialog contains the following information:

- Mapping ID:** observation-water
- Target (Triples Template):**

```

:observation/watersample/{mchimie}-{vmchimie_id} a oboe-core:Observation ;
obo-core:ofEntity :WaterSample ;
obo-core:hasMeasurement
:measurement/ph/watersample/{mchimie_id} ;
:measurement/o2/watersample/{mchimie_id} ;
:measurement/depth/watersample/{mchimie} ;
:measurement/depthmin/watersample/{mchimie} ;
:measurement/depthmax/watersample/{mchimie} ;
obo-core:hasContext :observation/plateforme/{loc_id} .

```
- Source (SQL Query):**

```

SELECT
mesure_chimie_mchimie.mchimie_id AS mchimie , valeur_mesure_chimie_vmchimie.vmchimie_id AS vmchimie_id, plateforme_pla.loc_id AS loc_id
FROM
mesure_chimie_mchimie
INNER JOIN valeur_mesure_chimie_vmchimie
ON mesure_chimie_mchimie.mchimie_id = valeur_mesure_chimie_vmchimie.mchimie_id
INNER JOIN sous_sequence_chimie_sschimie
ON sous_sequence_chimie_sschimie.sschimie_id = mesure_chimie_mchimie.sschimie_id
INNER JOIN plateforme_pla
ON sous_sequence_chimie_sschimie.loc_id = plateforme_pla.loc_id

```

At the bottom of the dialog, there are 'Update' and 'Cancel' buttons, and a 'Test SQL Query' checkbox. The background shows the 'Mapping manager' and 'Mapping Assistant' tabs, with 'observation-water' selected in the mapping list.

Publication des données

Choix des outils :

1. Serveur Jetty 8 (conteneur de servlet)
2. Triple store Sesame + console administration workbench
3. API Quest + (fichier .owl et .obda)

Publication des données


OpenRDF

Sesame server

Repositories

New repository

Delete repository

Explore

Summary

Namespaces

Contexts

Types

Explore

Query

Saved Queries

Export

Modify

SPARQL Update

Add

Remove

Clear

System

Information

Current Selections:

Sesame server: <http://si-ola.inra.fr:8181/openrdf-sesame> [change](#)

Repository: - none - [change](#)

User (optional): - none - [change](#)

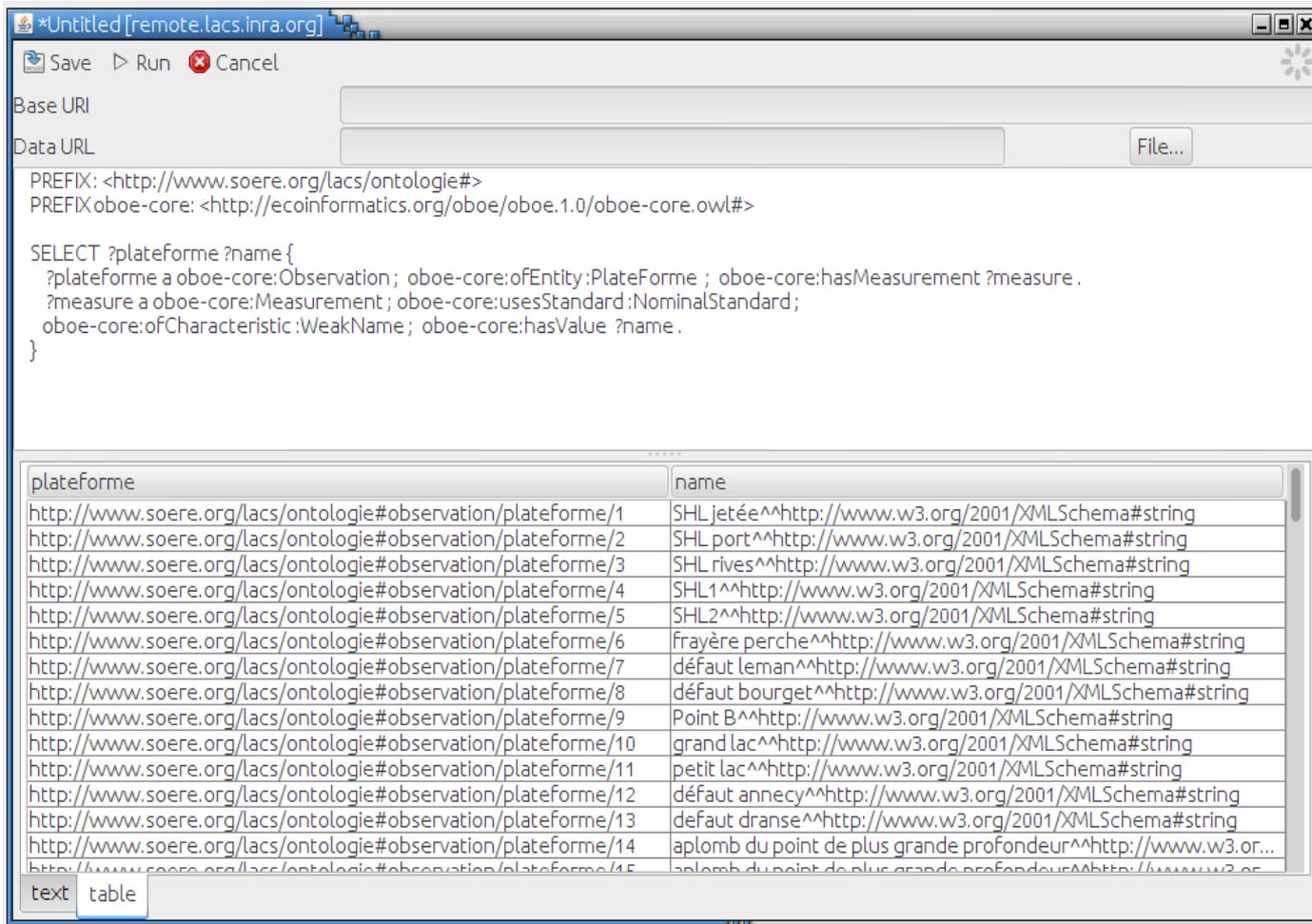
List of Repositories

		Id	Description	Location
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	SYSTEM	System configuration repository	http://si-ola.inra.fr:8181/openrdf-sesame/repositories/SYSTEM
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	lacs_mapping	Oboe-BD Lacs mapping	http://si-ola.inra.fr:8181/openrdf-sesame/repositories/lacs_mapping

Copyright © Aduna 1997-2011
Aduna - Semantic Power

Requêtage SPARQL (Twinkle)

End-point : <http://si-ola.inra.fr:8181..>



The screenshot shows the Twinkle SPARQL query editor. The query is as follows:

```

PREFIX: <http://www.soere.org/lacs/ontologie#>
PREFIX oboe-core: <http://ecoinformatics.org/oboe/oboe.1.0/oboe-core.owl#>

SELECT ?plateforme ?name {
  ?plateforme a oboe-core:Observation ; oboe-core:ofEntity:PlateForme ; oboe-core:hasMeasurement ?measure .
  ?measure a oboe-core:Measurement ; oboe-core:usesStandard:NominalStandard ;
  oboe-core:ofCharacteristic:WeakName ; oboe-core:hasValue ?name .
}
    
```

The results are displayed in a table with two columns: 'plateforme' and 'name'. The results are as follows:

plateforme	name
http://www.soere.org/lacs/ontologie#observation/plateforme/1	SHL jetée ^{^^} http://www.w3.org/2001/XMLSchema#string
http://www.soere.org/lacs/ontologie#observation/plateforme/2	SHL port ^{^^} http://www.w3.org/2001/XMLSchema#string
http://www.soere.org/lacs/ontologie#observation/plateforme/3	SHL rives ^{^^} http://www.w3.org/2001/XMLSchema#string
http://www.soere.org/lacs/ontologie#observation/plateforme/4	SHL1 ^{^^} http://www.w3.org/2001/XMLSchema#string
http://www.soere.org/lacs/ontologie#observation/plateforme/5	SHL2 ^{^^} http://www.w3.org/2001/XMLSchema#string
http://www.soere.org/lacs/ontologie#observation/plateforme/6	frayère perche ^{^^} http://www.w3.org/2001/XMLSchema#string
http://www.soere.org/lacs/ontologie#observation/plateforme/7	défaut leman ^{^^} http://www.w3.org/2001/XMLSchema#string
http://www.soere.org/lacs/ontologie#observation/plateforme/8	défaut bourget ^{^^} http://www.w3.org/2001/XMLSchema#string
http://www.soere.org/lacs/ontologie#observation/plateforme/9	Point B ^{^^} http://www.w3.org/2001/XMLSchema#string
http://www.soere.org/lacs/ontologie#observation/plateforme/10	grand lac ^{^^} http://www.w3.org/2001/XMLSchema#string
http://www.soere.org/lacs/ontologie#observation/plateforme/11	petit lac ^{^^} http://www.w3.org/2001/XMLSchema#string
http://www.soere.org/lacs/ontologie#observation/plateforme/12	défaut anecy ^{^^} http://www.w3.org/2001/XMLSchema#string
http://www.soere.org/lacs/ontologie#observation/plateforme/13	defaut dranse ^{^^} http://www.w3.org/2001/XMLSchema#string
http://www.soere.org/lacs/ontologie#observation/plateforme/14	aplomb du point de plus grande profondeur ^{^^} http://www.w3.org/2001/XMLSchema#string
http://www.soere.org/lacs/ontologie#observation/plateforme/15	aplomb du point de plus grande profondeur ^{^^} http://www.w3.org/2001/XMLSchema#string

Requêtage SPARQL (Twinkle)

End-point : <http://si-ola.inra.fr:8181..>

Save Run Cancel

Base URI

Data URL File...

```

PREFIX: <http://www.soere.org/lacs/ontologie#>
PREFIX oboe-core: <http://ecoinformatics.org/oboe/oboe.1.0/oboe-core.owl#>

SELECT ?obsWater ?measurPhUri ?valeurPH ?depth {

  ?obsWater a oboe-core:Observation ; oboe-core:ofEntity :WaterSample ; oboe-core:hasMeasurement ?measurPhUri , ?measurDepth .
  ?measurDepth oboe-core:ofCharacteristic :DepthAvg ; oboe-core:hasValue ?depth .
  ?measurPhUri oboe-core:ofCharacteristic :Ph ; oboe-core:hasValue ?valeurPH .
}

LIMIT 10000
      
```

obsWater	measurPhUri	valeurPH	depth
http://www.soere.org/lacs/ontologie#observation/watersample/1-7	http://www.soere.org/lacs/ontologie#measurement/ph/watersample/7	4.325E0^http://w...	0.0E0^http://ww...
http://www.soere.org/lacs/ontologie#observation/watersample/2-17	http://www.soere.org/lacs/ontologie#measurement/ph/watersample/17	4.305E0^http://w...	0.0E0^http://ww...
http://www.soere.org/lacs/ontologie#observation/watersample/3-27	http://www.soere.org/lacs/ontologie#measurement/ph/watersample/27	4.115E0^http://w...	0.0E0^http://ww...
http://www.soere.org/lacs/ontologie#observation/watersample/4-37	http://www.soere.org/lacs/ontologie#measurement/ph/watersample/37	4.24E0^http://w...	0.0E0^http://ww...
http://www.soere.org/lacs/ontologie#observation/watersample/5-47	http://www.soere.org/lacs/ontologie#measurement/ph/watersample/47	4.145E0^http://w...	0.0E0^http://ww...
http://www.soere.org/lacs/ontologie#observation/watersample/6-57	http://www.soere.org/lacs/ontologie#measurement/ph/watersample/57	4.145E0^http://w...	0.0E0^http://ww...
http://www.soere.org/lacs/ontologie#observation/watersample/7-67	http://www.soere.org/lacs/ontologie#measurement/ph/watersample/67	4.315E0^http://w...	0.0E0^http://ww...
http://www.soere.org/lacs/ontologie#observation/watersample/8-77	http://www.soere.org/lacs/ontologie#measurement/ph/watersample/77	4.065E0^http://w...	0.0E0^http://ww...
http://www.soere.org/lacs/ontologie#observation/watersample/9-87	http://www.soere.org/lacs/ontologie#measurement/ph/watersample/87	4.31E0^http://w...	0.0E0^http://ww...
http://www.soere.org/lacs/ontologie#observation/watersample/10-97	http://www.soere.org/lacs/ontologie#measurement/ph/watersample/97	4.31E0^http://w...	0.0E0^http://ww...
http://www.soere.org/lacs/ontologie#observation/watersample/11-107	http://www.soere.org/lacs/ontologie#measurement/ph/watersample/107	4.175E0^http://w...	0.0E0^http://ww...
http://www.soere.org/lacs/ontologie#observation/watersample/12-117	http://www.soere.org/lacs/ontologie#measurement/ph/watersample/117	4.155E0^http://w...	0.0E0^http://ww...
http://www.soere.org/lacs/ontologie#observation/watersample/13-127	http://www.soere.org/lacs/ontologie#measurement/ph/watersample/127	4.135E0^http://w...	0.0E0^http://ww...

text table

Aller plus loin...

- Gestion de droits des utilisateurs
- Contrôle du volume de données renvoyées
- Augmenter les performances via du traitement //

Pour cela :

Au lieu d'exposer le SPARQL end-point directement sur le web, exposer un web service (REST ou SOAP) qui se chargera d'intercepter les requêtes SPARQL des utilisateurs et appeler le SPARQL end-point en local.

MERCI DE VOTRE ATTENTION