

➤ Ontologie ... ontologies du Web

Catherine Roussey, équipe Copain, UR
TSCF, Centre Clermont ARA

Plan

Definitions: connaissances, ontologie

Technologies du Web sémantique → LOD

Publication de données capteurs sur le LOD

Architecture d'intégration de données

Perspectives



Définitions

“Une donnée est une information qui peut être enregistrée, traitée, analysée ou communiquée, quelle que soit sa nature. Lorsque la donnée n’a pas encore été traitée ou contextualisée, on parle de donnée brute”. (Guide pratique pour la gestion des données de la recherche irstea)

Donnée: un élément d’information,
percevable,
manipulable

Information: donnée +
sens + contexte
Type → structure de données

Connaissance: information formalisée dans un langage...
Pour inférer de nouvelles connaissances
“marié à ” est une relation symétrique

Ontologie: un peu d'histoire

Il était une fois, les systèmes experts une branche de l'Intelligence Artificielle dans les années 90, évolution vers les systèmes à base de connaissances
Séparer les connaissances de résolution de problème des connaissances du domaine

Ontologie: formalisation des connaissances du domaine consensuelles et stables

Ontologie: *“spécification explicite d'une conceptualisation”* (Gruber, 1993)

Problème: spécification + conceptualisation = diagramme UML

Ontologie *“une spécification explicite et formelle d'une conceptualisation partagée”* (Studer et al., 1998)

Ontologie du web: les schémas de données formalisés dans un des langages des technologies du Web sémantique

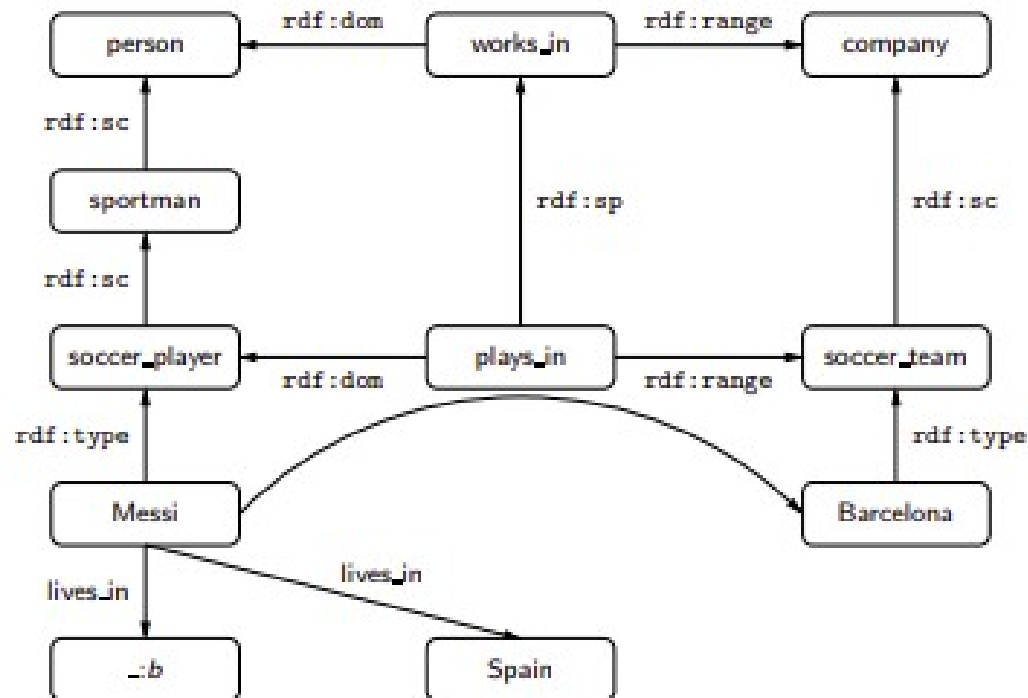
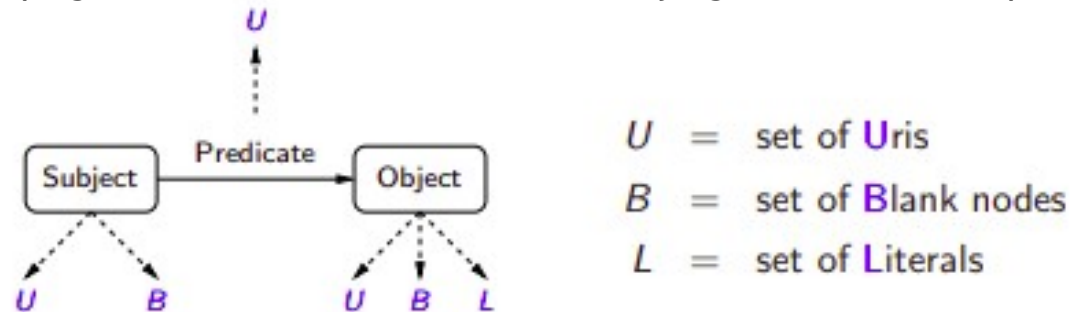


Resource Description Framework RDF

Credit Semantic Web Data/RDF/SPARQL

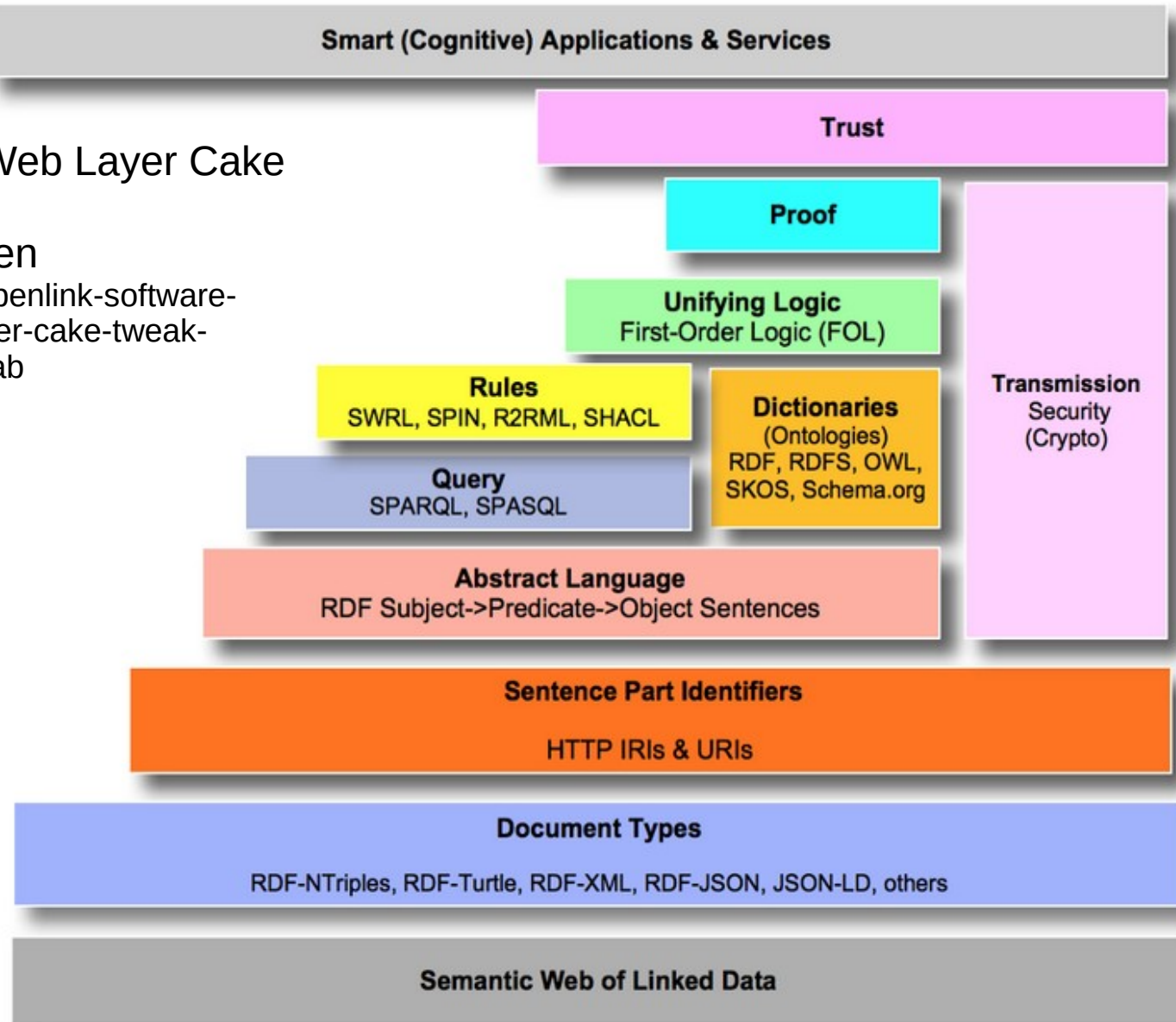
<http://homepages.inf.ed.ac.uk/libkin/teach/beijing2018/rdfslides.pdf>

Triplet RDF



Technologies Web Sémantique

Credit Semantic Web Layer Cake
Tweak, Explained
Kingsley Uyi Idehen
<https://medium.com/openlink-software-blog/semantic-web-layer-cake-tweak-explained-6ba5c6ac3fab>



Ontologie du Web W3C

Dans le monde du web sémantique, une ontologie est l'ensemble des concepts et des relations utilisés pour décrire un domaine d'intérêt. Les mots ontologie et vocabulaire sont utilisés conjointement. Le mot ontologie est employé quand le vocabulaire de concepts et de relations est assez complexes et peut contenir par exemple des contraintes: conditions nécessaire et/ou suffisante d'appartenance. (W3C)

Une ontologie sert à (W3C):

- **Normaliser** les termes du domaine: leur associer un identifiant (URI), un label et une signification
- **Typer** les éléments de ce vocabulaire pour définir un schéma, documenté et réutilisable: des classes, des propriétés etc...
- Aider l'intégration de données multi-sources
- Organiser les connaissances d'un domaine: publication de ressources et de leur métadonnées descriptives sur le web (Linked Data)
- Produire des **inférences**



Linked Open Data (LOD)

Web de données Liées

An extension of the current Web...

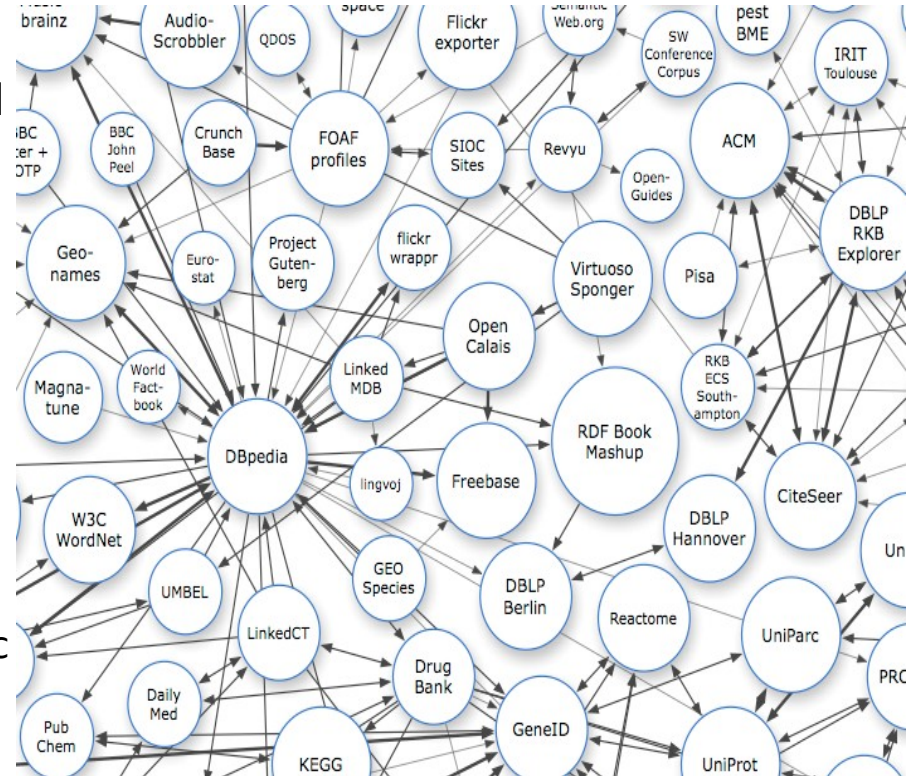
... where data are given well-defined and explicitly represented meaning, ...

... so that it can be shared and used by humans and machines, ...

... better enabling them to work in cooperation

And clear principles on how to publish data

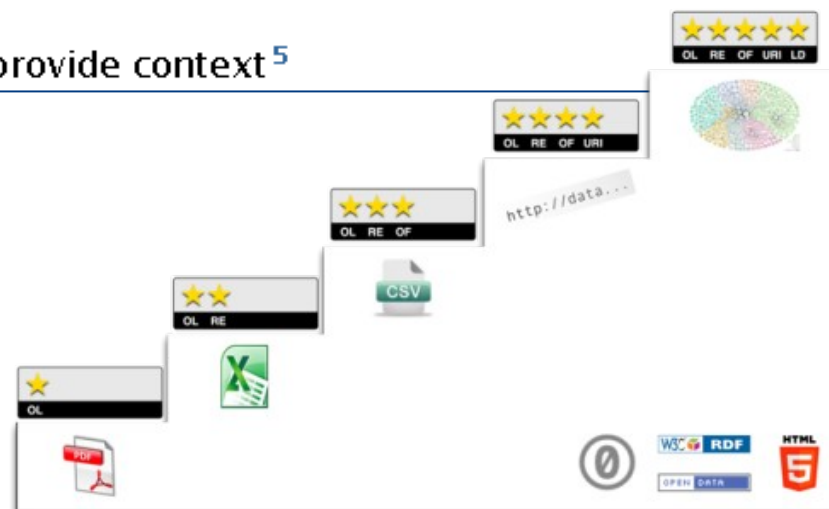
Credit: Oscar Corcho "Ingredients for Semantic Sensor Networks"



Modèle 5 étoiles

Credit Tim Berners-Lee, <http://5stardata.info/>

- ★ make your stuff available on the Web (whatever format) under an open license¹
- ★★ make it available as structured data (e.g., Excel instead of image scan of a table)²
- ★★★ use non-proprietary formats (e.g., CSV instead of Excel)³
- ★★★★ use URIs to identify things, so that people can point at your stuff⁴
- ★★★★★ link your data to other data to provide context⁵



Données FAIR



Credit FAIR in a nutshell. Image: ARDC 2018 - CC-BY 4.0.

<https://librarycarpentry.org/Top-10-FAIR/2018/12/01/aus-government-data/>

INRAE

Ontologies du Web

Octobre 2020 / C Roussey

INRAE fournisseur de données

INRAE possède :

- des réseaux de capteurs pour observer une variété de phénomènes naturels: météo, hydro, pollution etc...
- des bases relationnelles pour stocker les données d'expérimentation, ...
- des bases bibliographiques: bulletins d'alertes, articles scientifiques, ...

Comment rendre accessible ces données ?

au personnel d'INRAE mais pas que...

Le site web BDOH permet de rechercher les fichiers de données capteurs (séries temporelles) des observatoires hydrologiques d'INRAE (ex-Irstea)



Hétérogénéité ...

Données de capteurs: observations automatisées qui doivent être structurées, complétées, validées et documentées.

Différents types de capteurs: observation d'une variété de phénomènes naturels

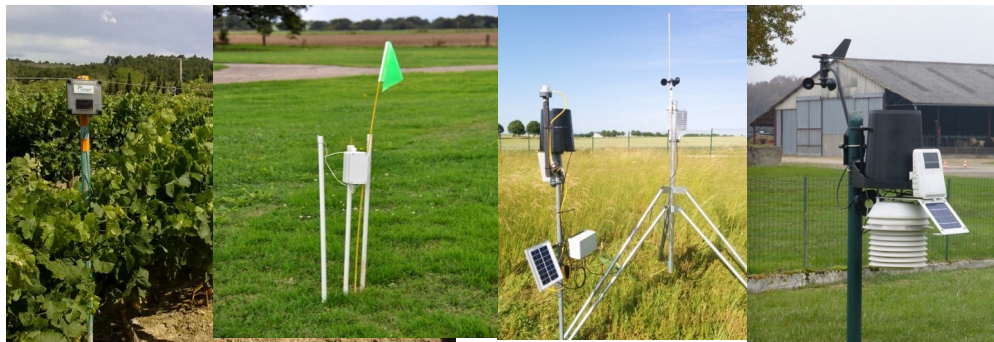
Différentes types de données: des mesures aux métadonnées descriptives des capteurs

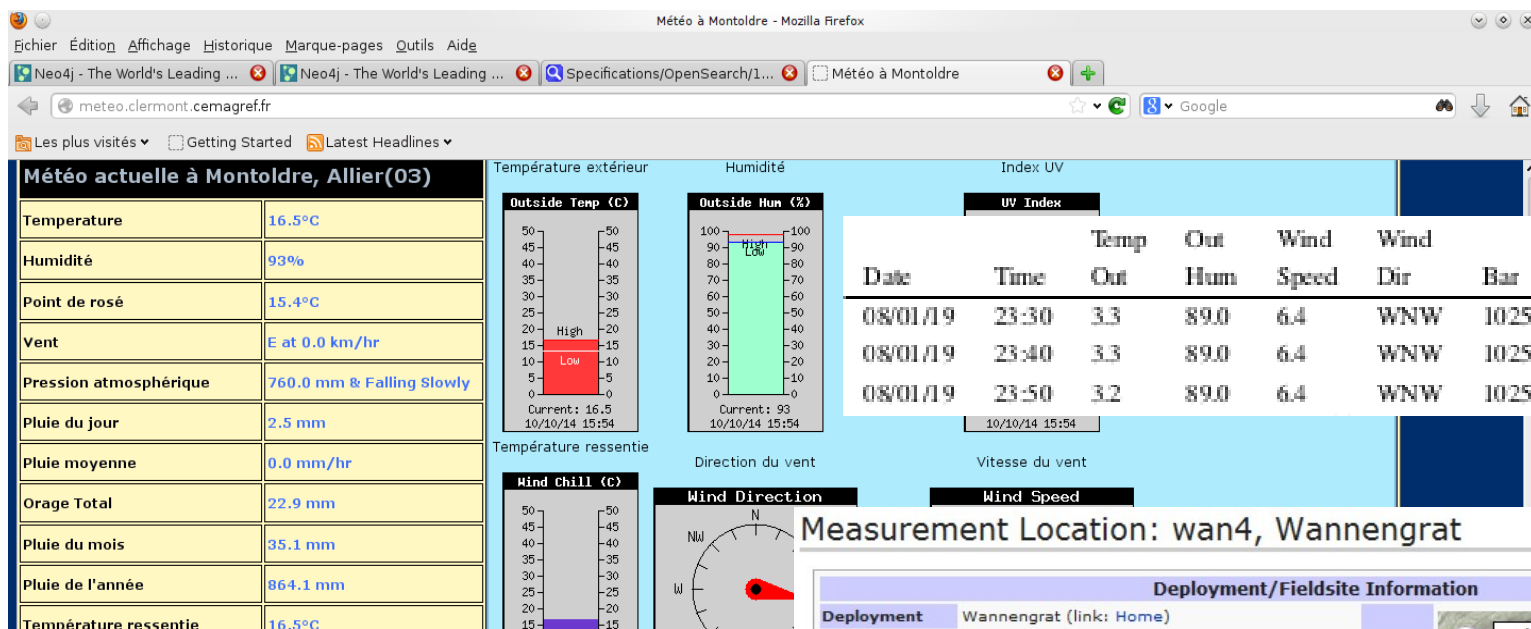
Différents formats / schéma de données / interface d'interrogation

Différents usages: agriculture, gestion des risques, etc...

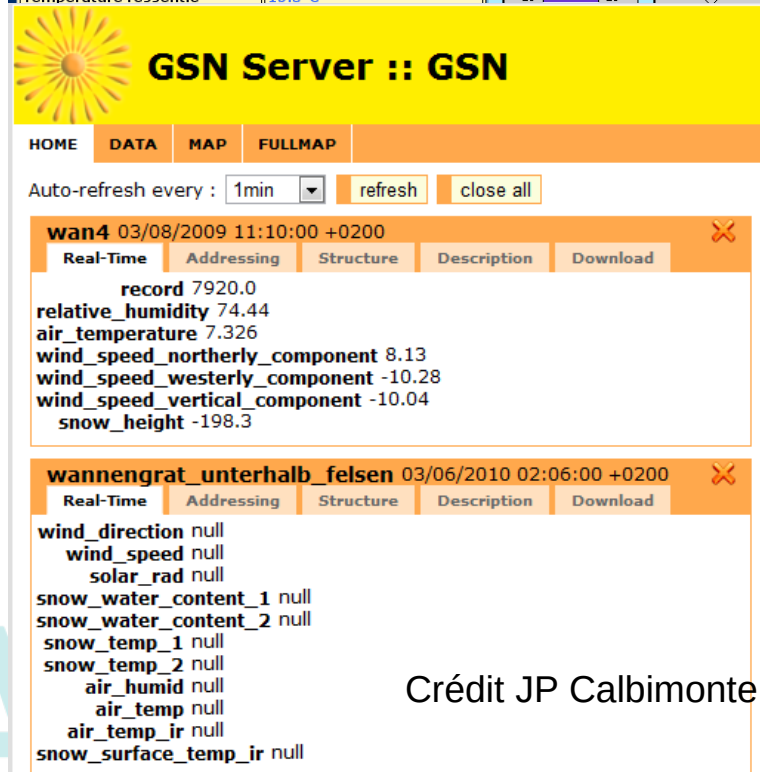
→ Difficulté à interroger de manière unifiée les données issues de différentes stations

→ Difficulté à intégrer des données de divers réseaux dans une même application



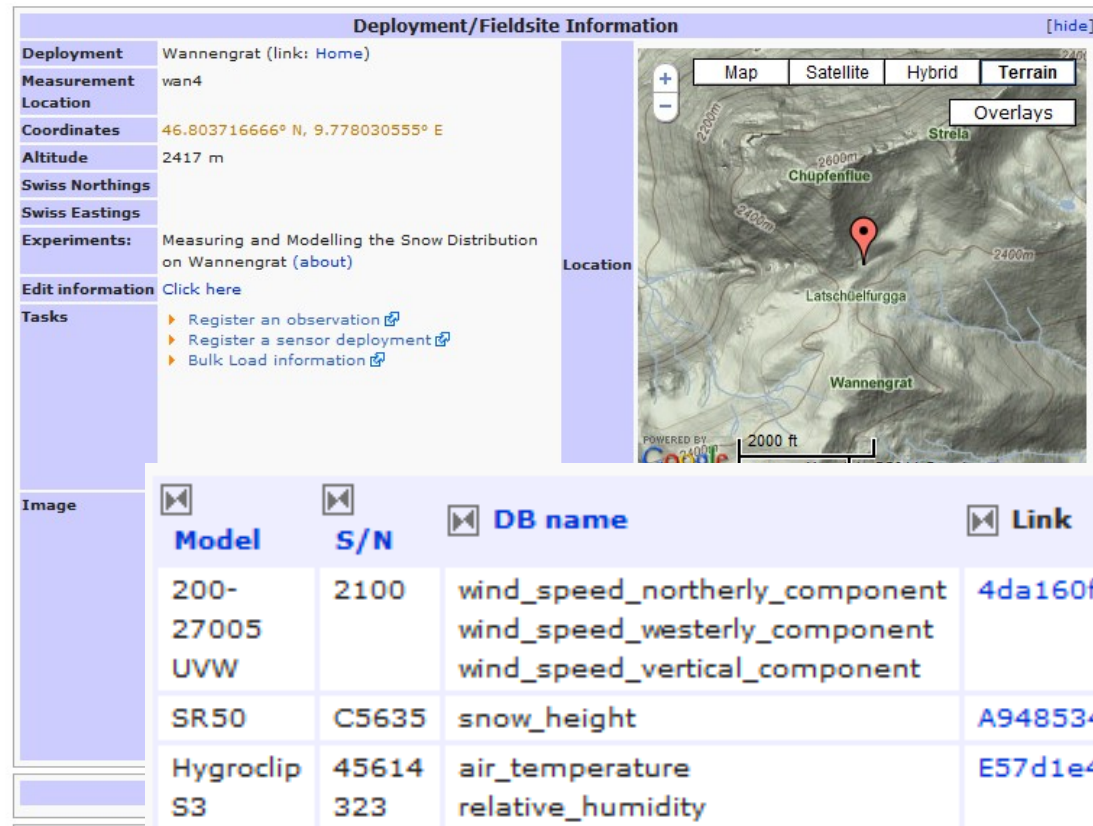


Crédit G André



Crédit JP Calbimonte

Measurement Location: wan4, Wannengrat



Publication de données capteurs

Accès uniformisé aux données et métadonnées des capteurs

« Standardisation » des données:

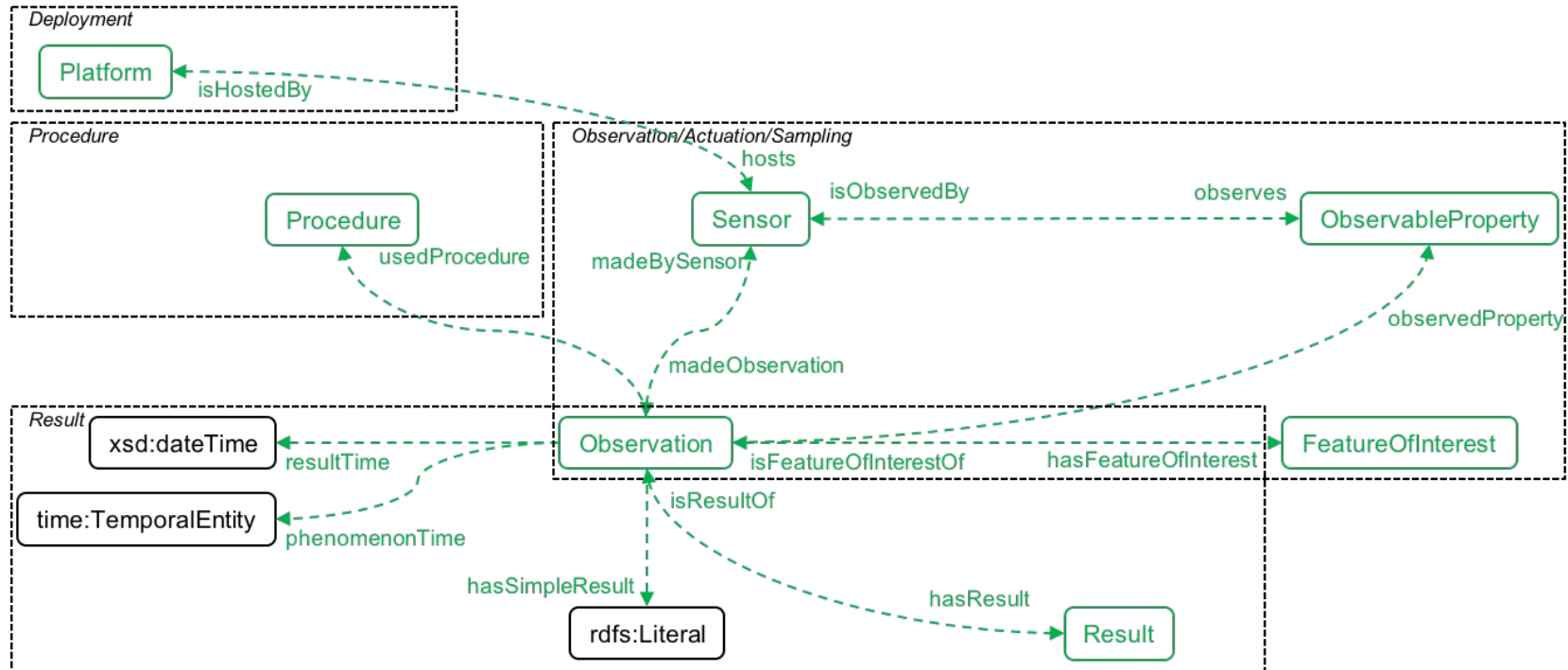
- Réutilisation de schémas (ontologies) existants déjà documentés
- Combinaison de différents schémas pour coller à notre besoin
- Possibilité d'étendre les schémas en fonction des besoins
- Facilité l'intégration de nos données avec d'autres jeux de données
- Facilité la découverte de notre jeu de données (visibilité)

Credit C. ROUSSEY, S. BERNARD, G. ANDRÉ, D. BOFFETY. Weather Data Publication on the LOD using SOSA/SSN Ontology. Semantic Web Journal, 2019
<http://www.semantic-web-journal.net/content/weather-data-publication-lod-using-sosassn-ontology-0>



Semantic Sensor Network (SSN)

Sensor Observation Sample Actuator (SOSA)



Credit W3C working group
<https://www.w3.org/TR/vocab-ssn/>



INRAE

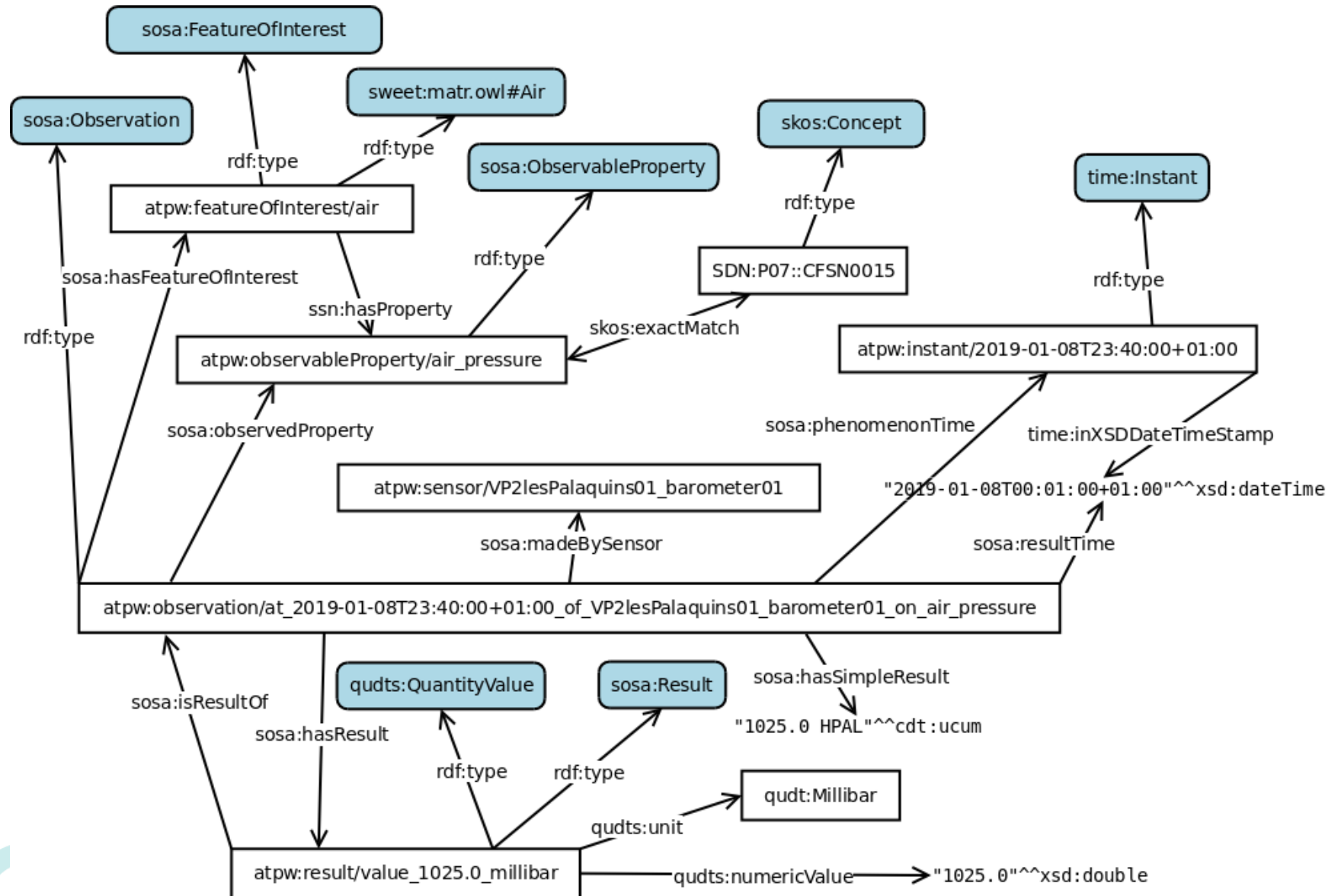
Ontologies du Web
Octobre 2020 / C Roussey

Réseau d'Ontologies

Semantic Sensor Network : le squelette

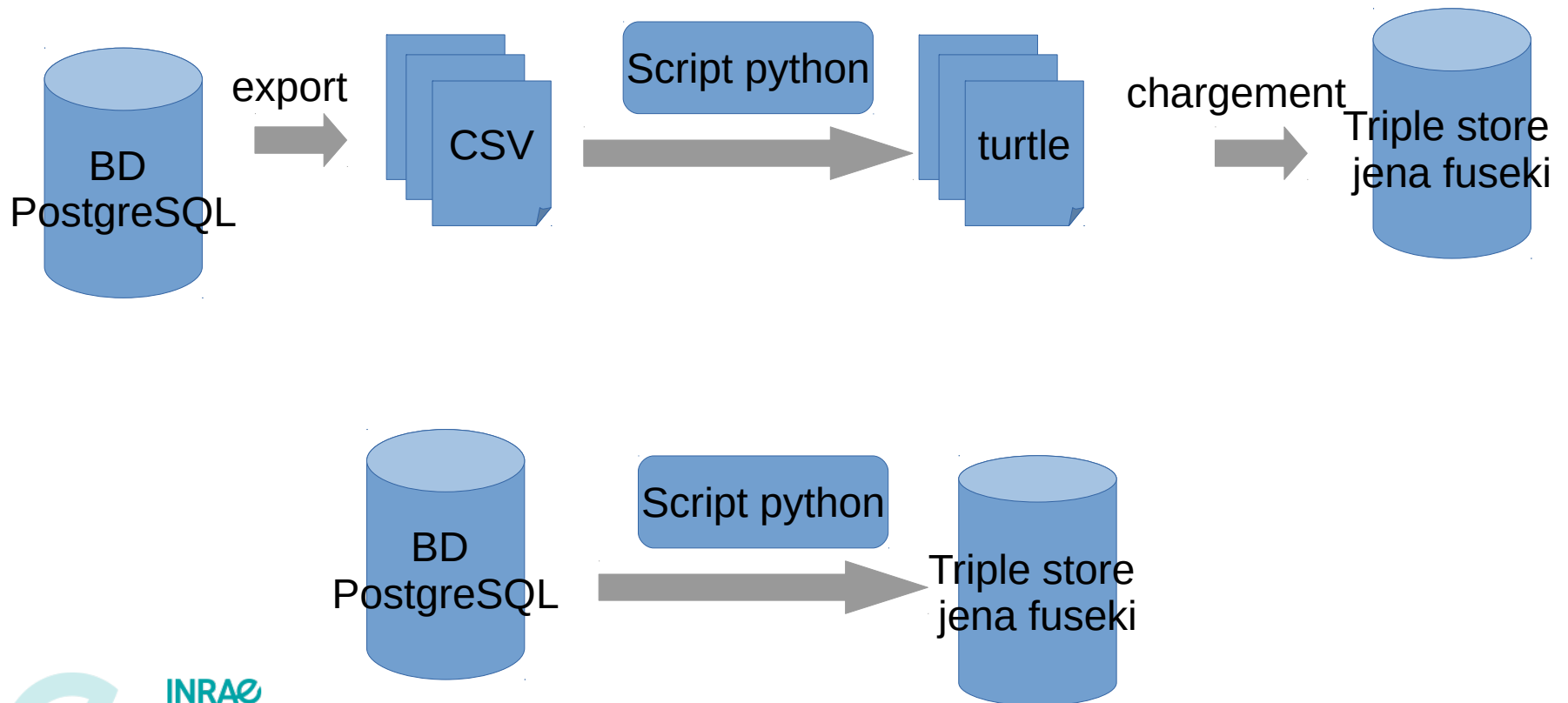
- Sensing Device
Ontology for meteorological sensor (aws)
- Feature of Interest
SWEET: set of classes
- Observable Property
Thesaurus Climate and Forecast (skos)
- Platform location
GeoSPARQL and Location Core Vocabulary (geosparql + locn)
- Observation result and phenomenon time
W3C Time Ontology (time)
- Observation value
Library of Quantity Kind and Units (qu + dim)
Dolce Ultra Light (dul)

Observation



BD to triple store

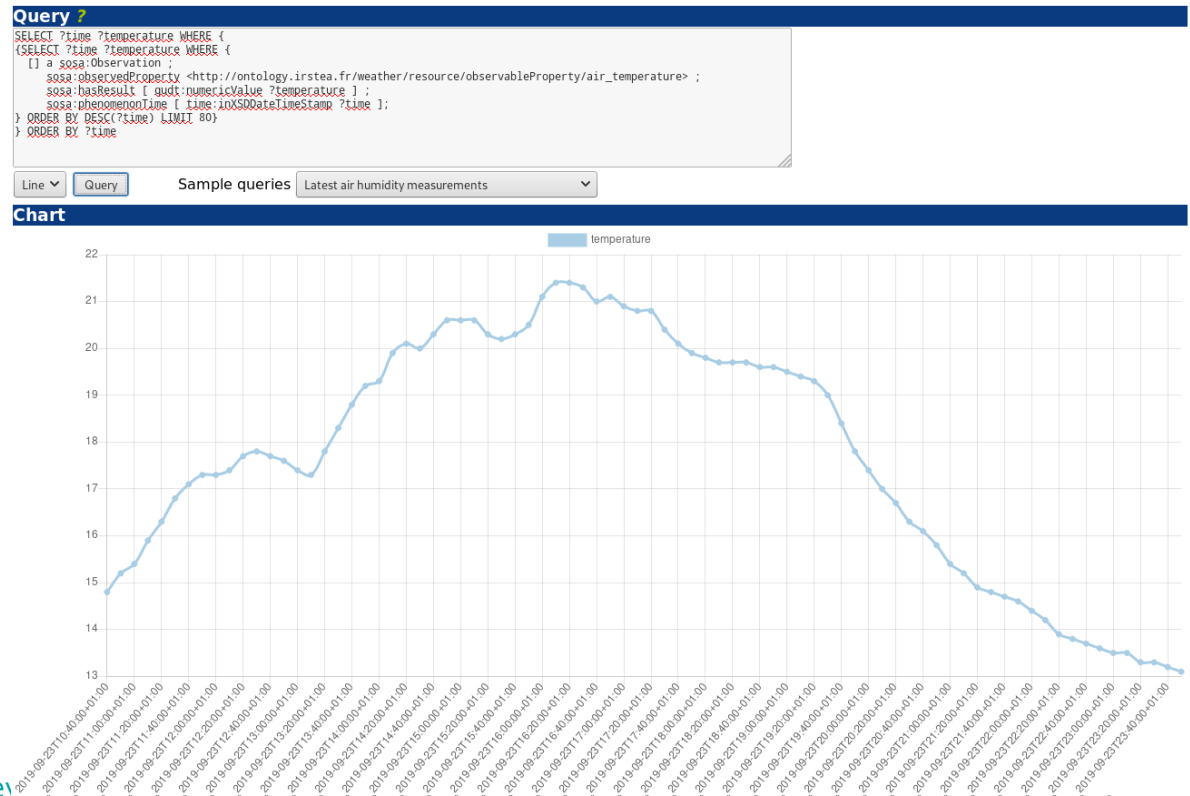
Credit Stephan Bernard stephan.bernard@inrae.fr



Interface d'interrogation

A SPARQL query interface using snorql is available at <http://ontology.irstea.fr/weather/snorql/>.

A SPARQL query interface that display numeric results as a graph is available at <http://ontology.irstea.fr/weather/snorql/test/>



INRAE

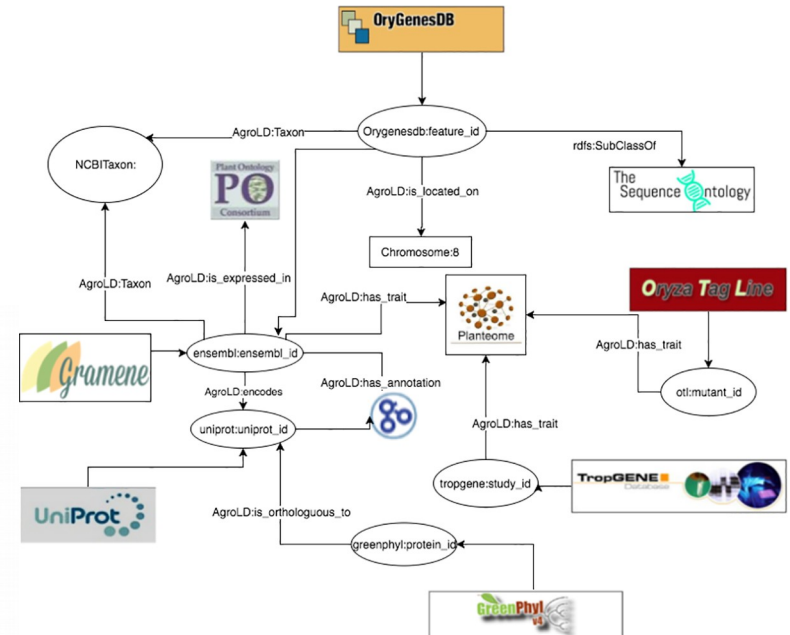
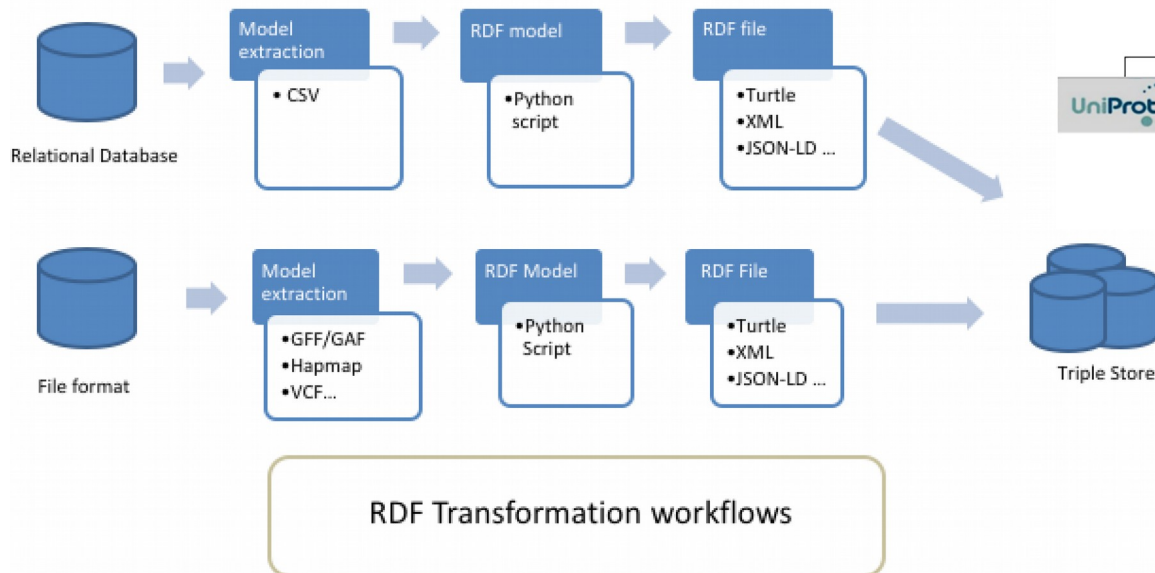
Ontologies du Web

Octobre 2020 / C Roussey

Agronomic Linked Data (AgroLD)

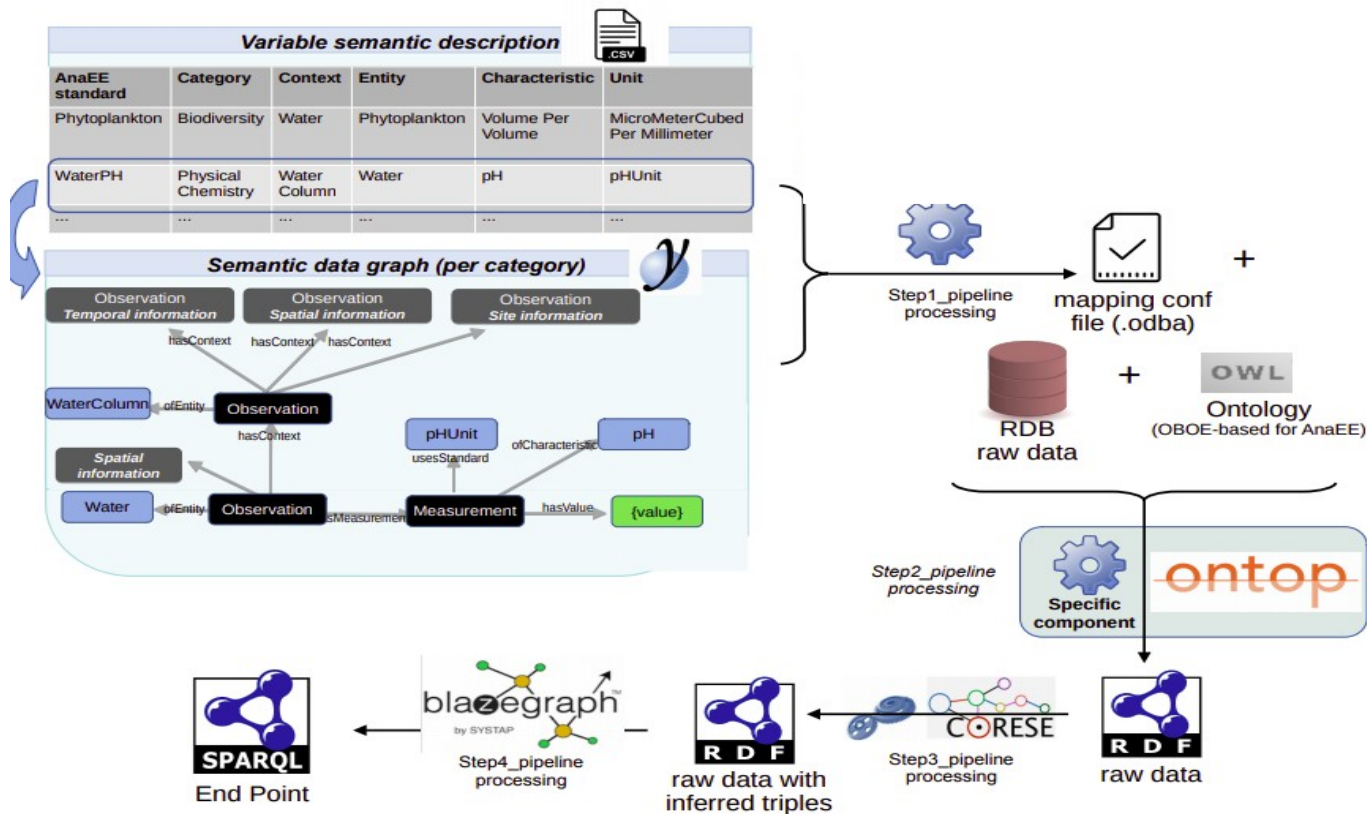
A knowledge-based system to enable integrative biology in agronomy

ETL process



Credit <https://doi.org/10.1371/journal.pone.0198270>

Pipelines for interoperability of distributed resources in ecology: the AnaEE-France context



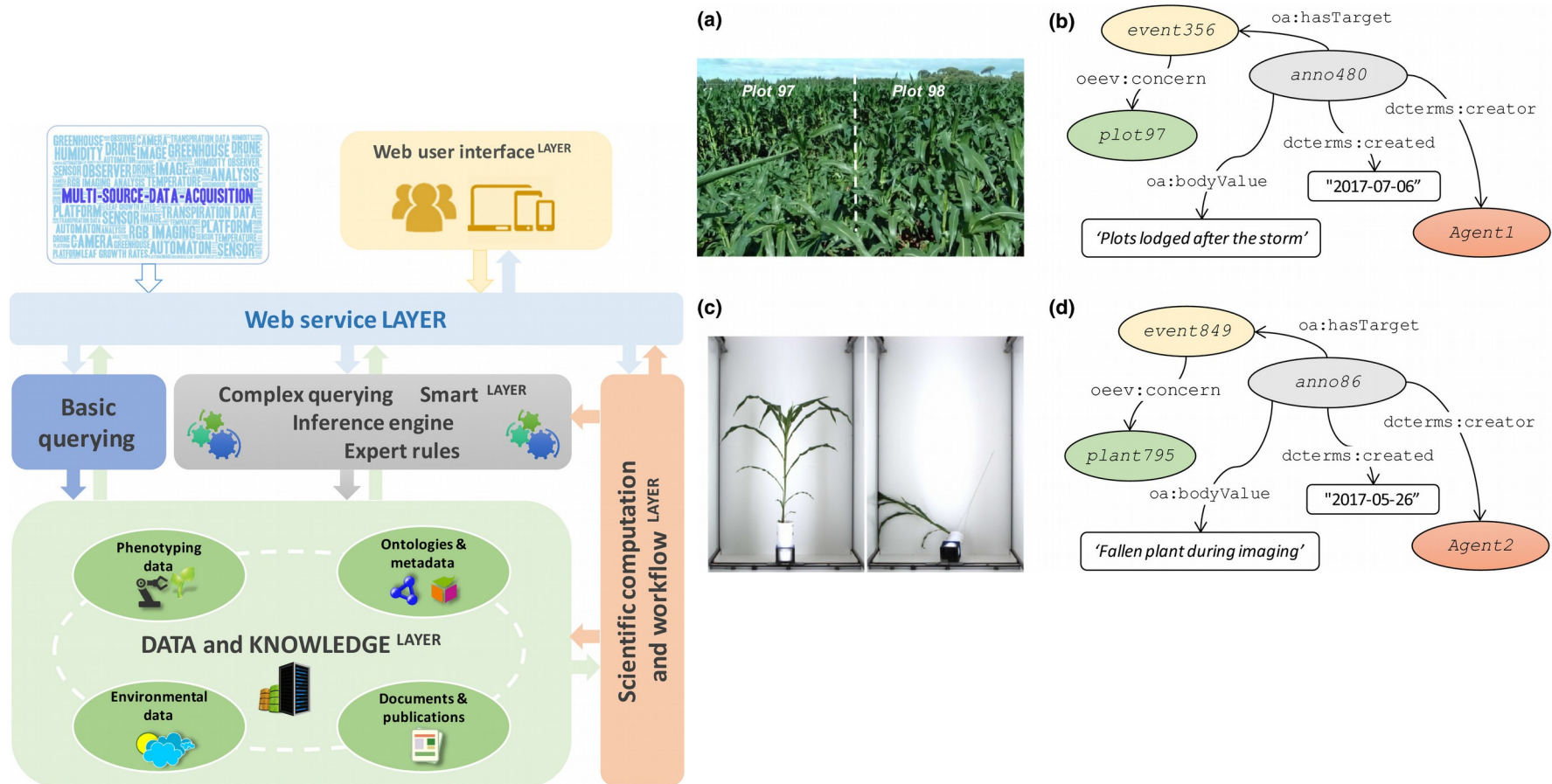
Credit Christian Pichot, Philippe Clastre, Benjamin Jaillet, Damien Maurice, Rachid Yahiaoui. 2020. 5ème édition des Journées Nationales du Développement Logiciel (JDEV2020), 7-10 Jul 2020, Rennes, France

http://devlog.cnrs.fr/_media/jdev2020/poster_jdev2020_t5_semanticinteroperabilitypipelines_cpichot_v2.pdf

Architecture OPEN SILEX:

Credit Dealing with multi-source and multi-scale information in plant phenomics: the ontology-driven Phenotyping Hybrid Information System

New Phytologist, Volume: 221, Issue: 1, Pages: 588-601, First published: 28 August 2018, DOI: (10.1111/nph.15385)



Perspectives

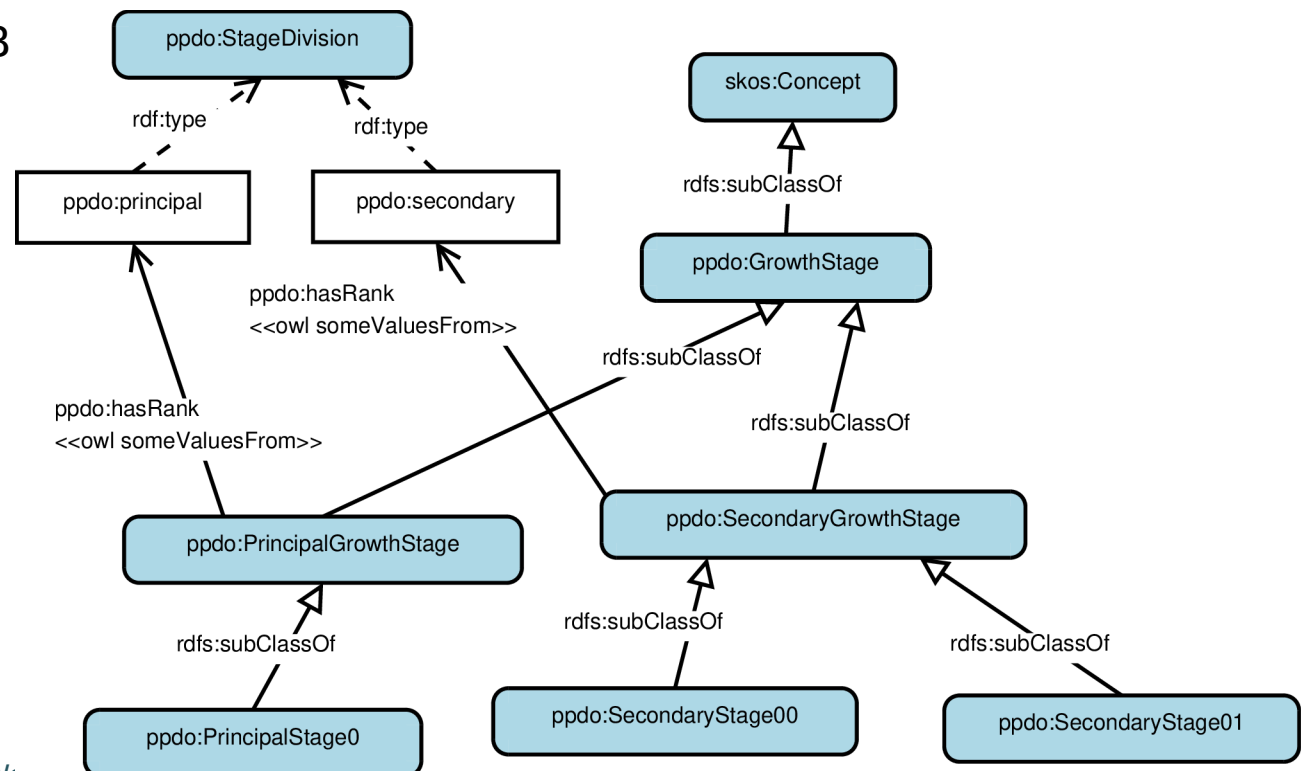
Annotation de BD + Annotation de texte

Interrogation multisources: BD d'épidémie versus bulletins d'alerte

Mieux intégrer les modèles SKOS dans les réseaux d'ontologies

SKOS pour publier des données de référence comme les stades phénologiques

Projet ANR D2KAB



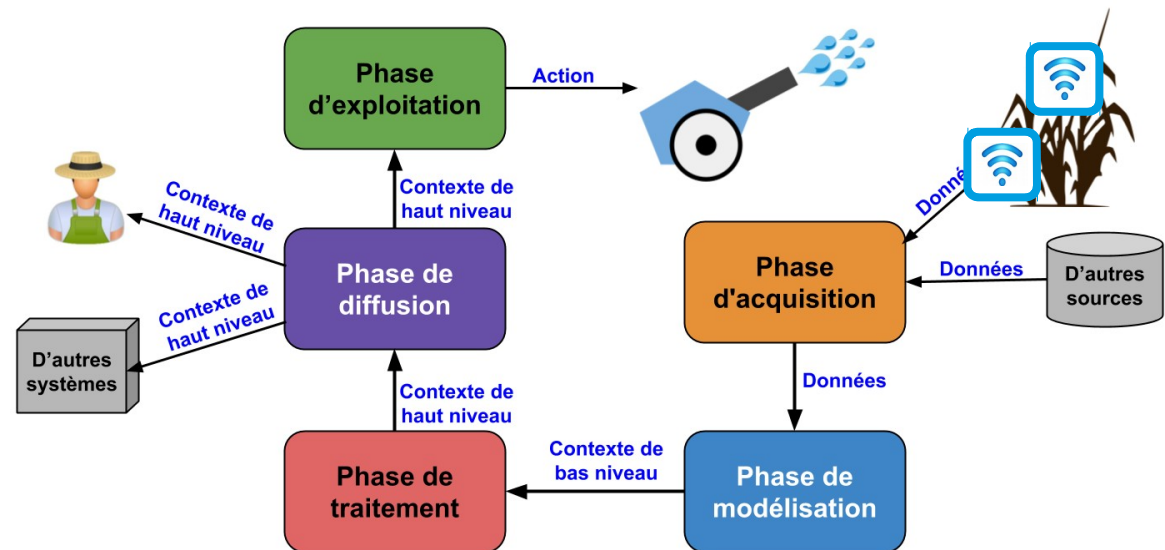
Perspectives

Système contextuel adaptatif

Un système contextuel est un système qui utilise le contexte pour fournir des informations et des services appropriés à l'utilisateur. (Abowd et al., 1999)

Implanter la boucle observation → décision → action → observation

Projet ANR COSWOT



Credit Q-D. NGUYEN, C. ROUSSEY, M. POVEDA-VILLALÓN, C. DE VAULX, J-P. CHANET. Development Experience of a Context-Aware System for Smart Irrigation Using CASO and IRRIG Ontologies. Applied Science 2020, 10(5), 1803; <https://doi.org/10.3390/app10051803>

Perspectives

Combiner ontologie d'observation + ontologie de service

SSN décrit une situation de mesure : qui, quoi, quand, comment.

Smart Appliance Reference (SAREF) décrit des appareils connectés de tous les domaines.

des ontologies noyaux auxquelles d'autres ontologies se connectent pour définir le schéma de données de l'application cible.

Ces deux ontologies s'appliquent à des domaines variés: agriculture, santé, domotique.

