



Introduction à la formalisation des connaissances sous forme d'ontologies

Catherine Roussey, Stéphan Bernard, Ghislain Atemezing

► To cite this version:

Catherine Roussey, Stéphan Bernard, Ghislain Atemezing. Introduction à la formalisation des connaissances sous forme d'ontologies. Master. BiSemWeb, Université de Labé République de Guinée, Guinée. 2023, pp.59. hal-04289780

HAL Id: hal-04289780

<https://hal.science/hal-04289780>

Submitted on 16 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Introduction à la ➤ formalisation des connaissances sous forme d'ontologies

Catherine ROUSSEY, Stephan BERNARD, Ghislain ATEMEZING

équipe Copain, UR TSCF, INRAE Centre Clermont ARA

Axe Informatique, UMR MISTEA, INRAE Centre Occitanie Montpellier

European Union Agency for Railways, Valenciennes



Plan

Definitions: connaissances, ontologie

Technologies du Web sémantique, graphe de connaissances → LOD

Ingénierie des ontologies

3 cas d'usages

1) Publication de données capteurs sur le LOD

2) Projet COSWOT : Système contextuel adaptatif

3) Projet D2KAB : Annotation sémantique de bulletins d'alerte agricole



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)

34

Donnée: un élément d’information,

percevable par l’humain,



manipulable par la machine



Information: donnée + sens

Sens : il faut fournir à l’humain du contexte



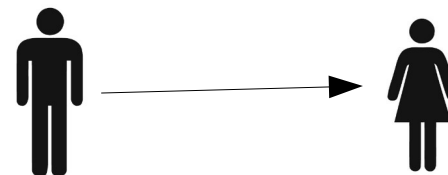
Type → structure de données



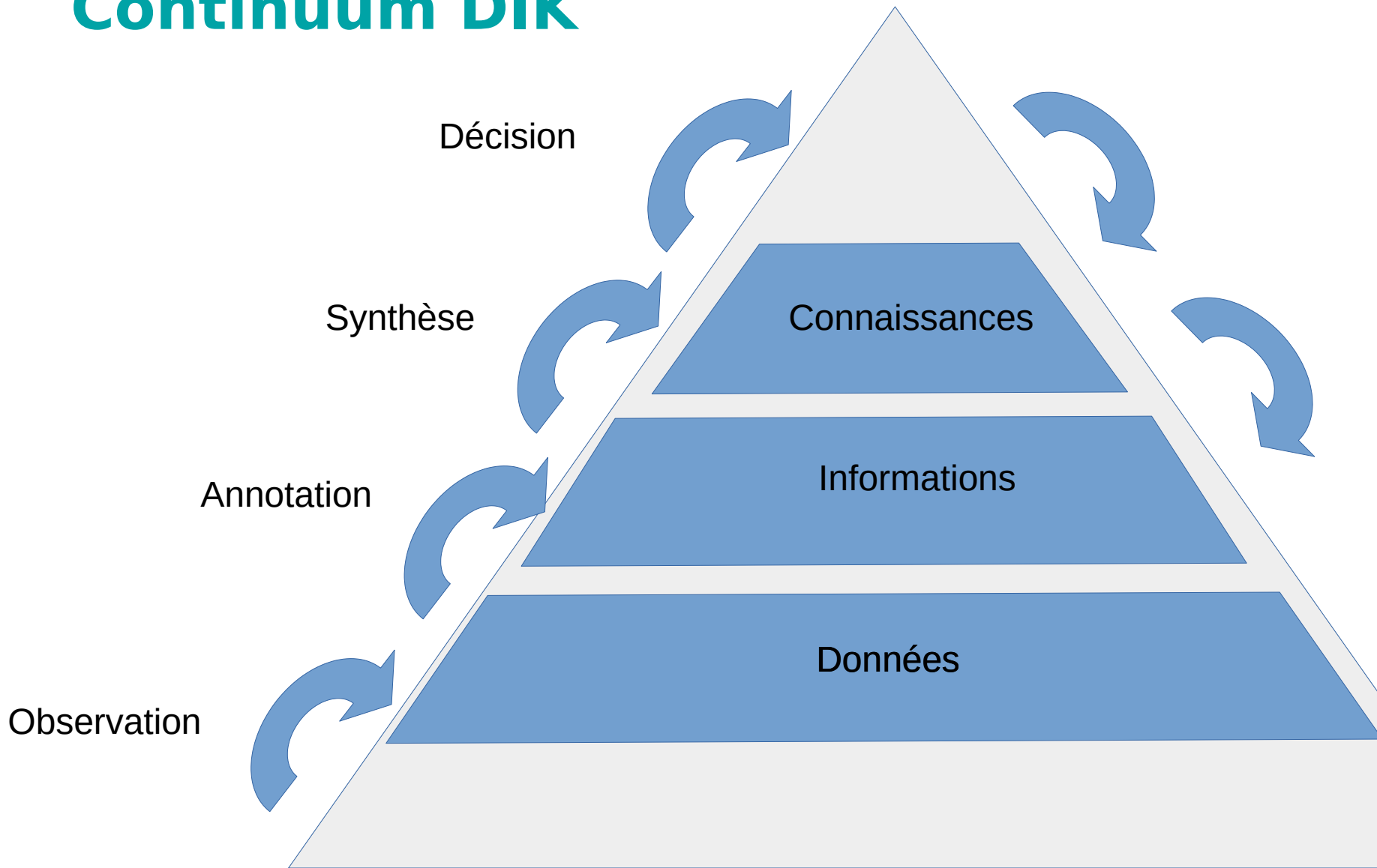
Connaissance: information formalisée dans un langage...

Pour inférer de nouvelles connaissances

“marié à ” est une relation symétrique



Continuum DIK



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

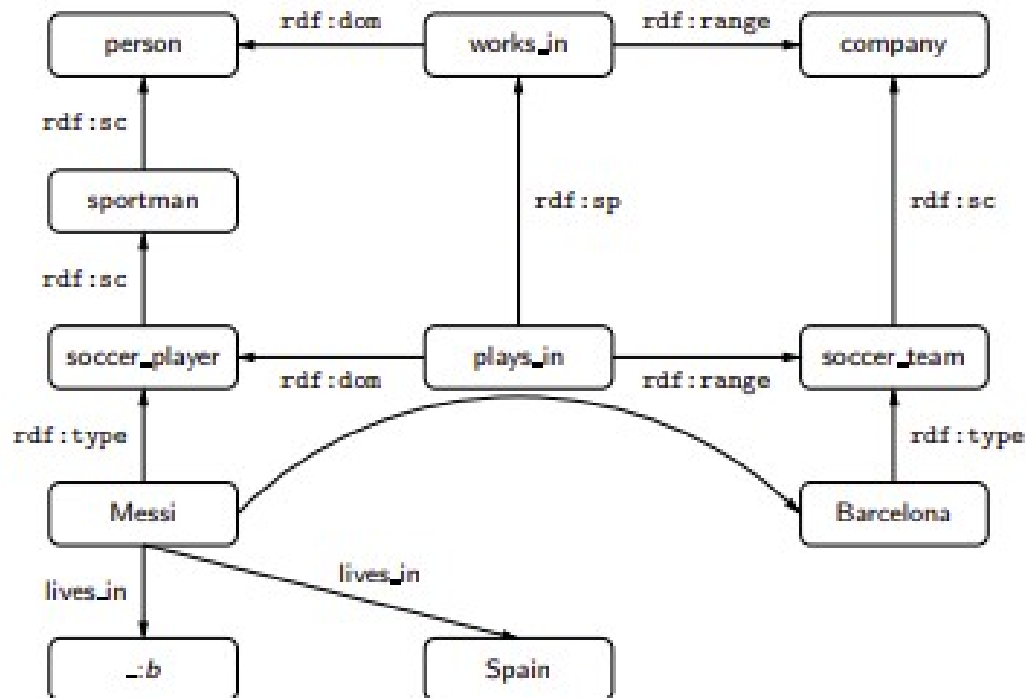
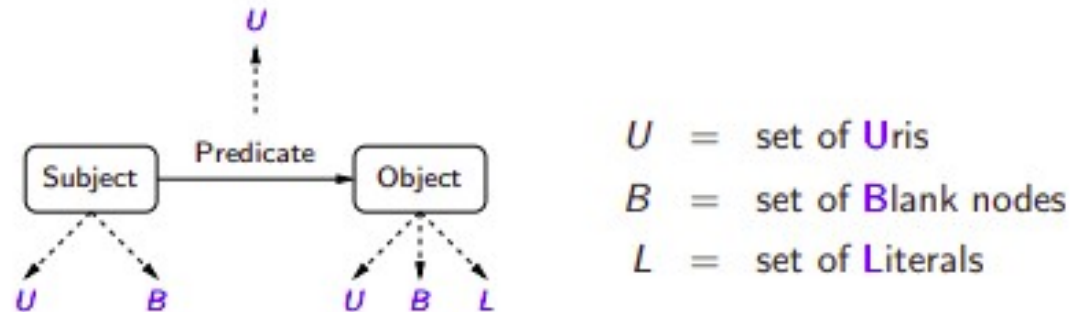


Resource Description Framework RDF

Credit Semantic Web Data/RDF/SPARQL

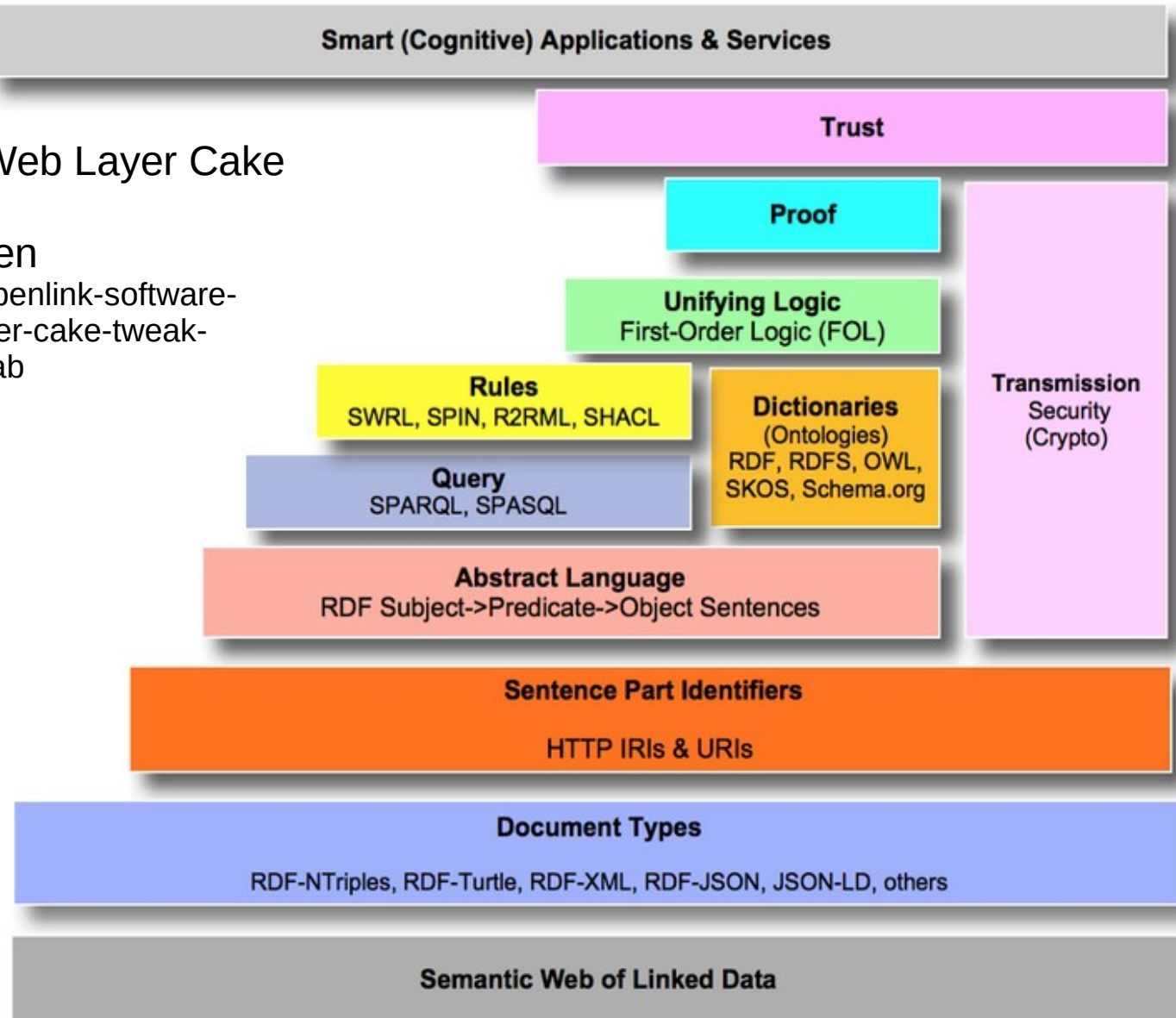
<http://homepages.inf.ed.ac.uk/libkin/teach/beijing2018/rdfslides.pdf>

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

et graphe de connaissances

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**

Un graphe de connaissances est l'ensemble des faits décrit à l'aide d'une ontologie et publié sur le web



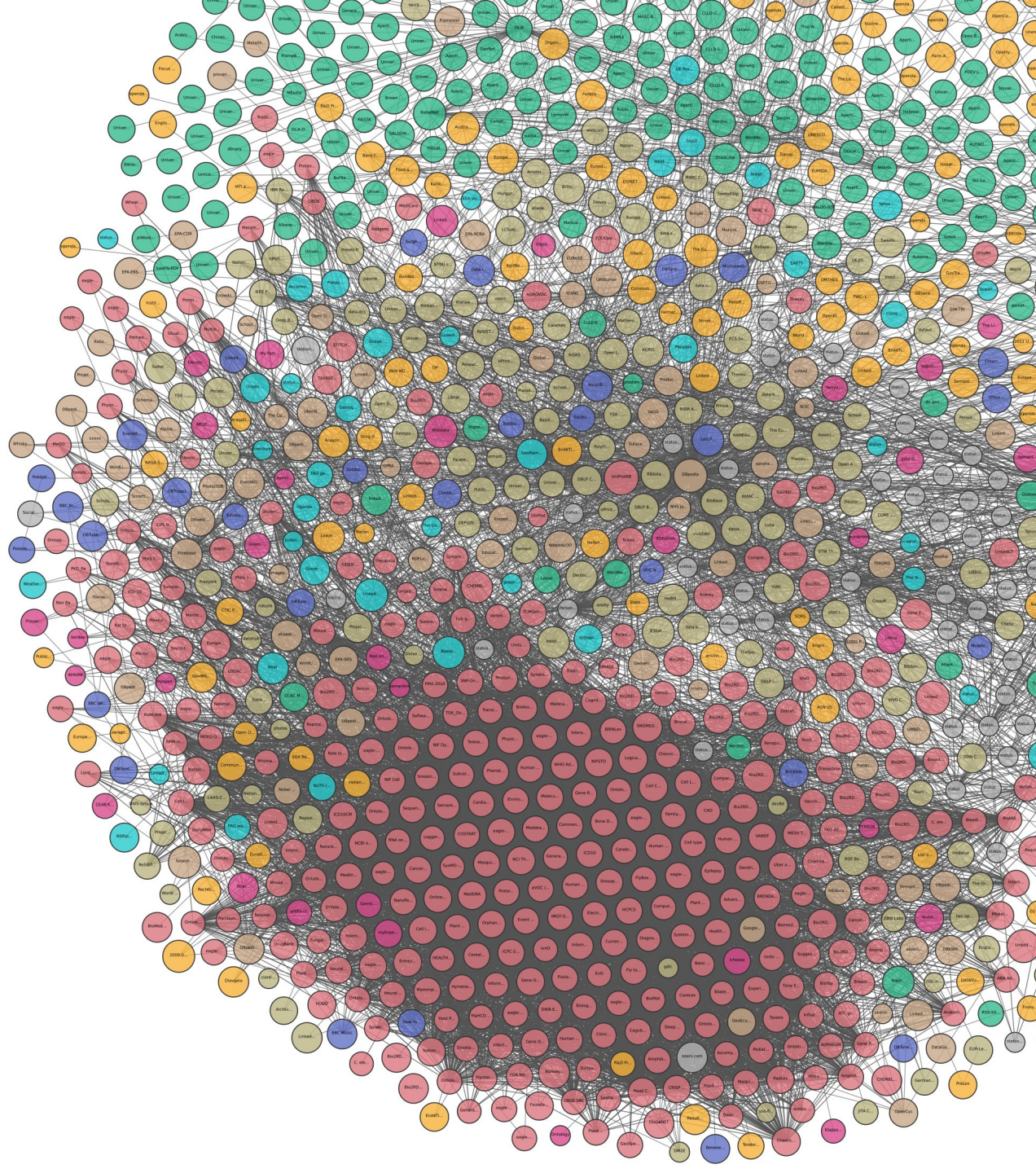
Linked Open Data (LOD)

Credit lod-cloud.net

Web de données Liées

Video europeana qu'est ce que le web des données

<https://vimeo.com/49231111>



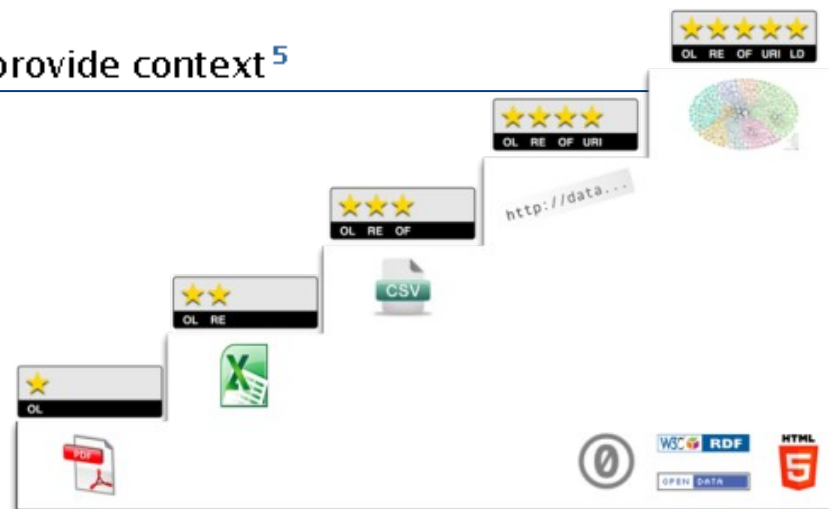
INRAE

Formalisation des connaissances
Novembre 2023 / C Rousse

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/>

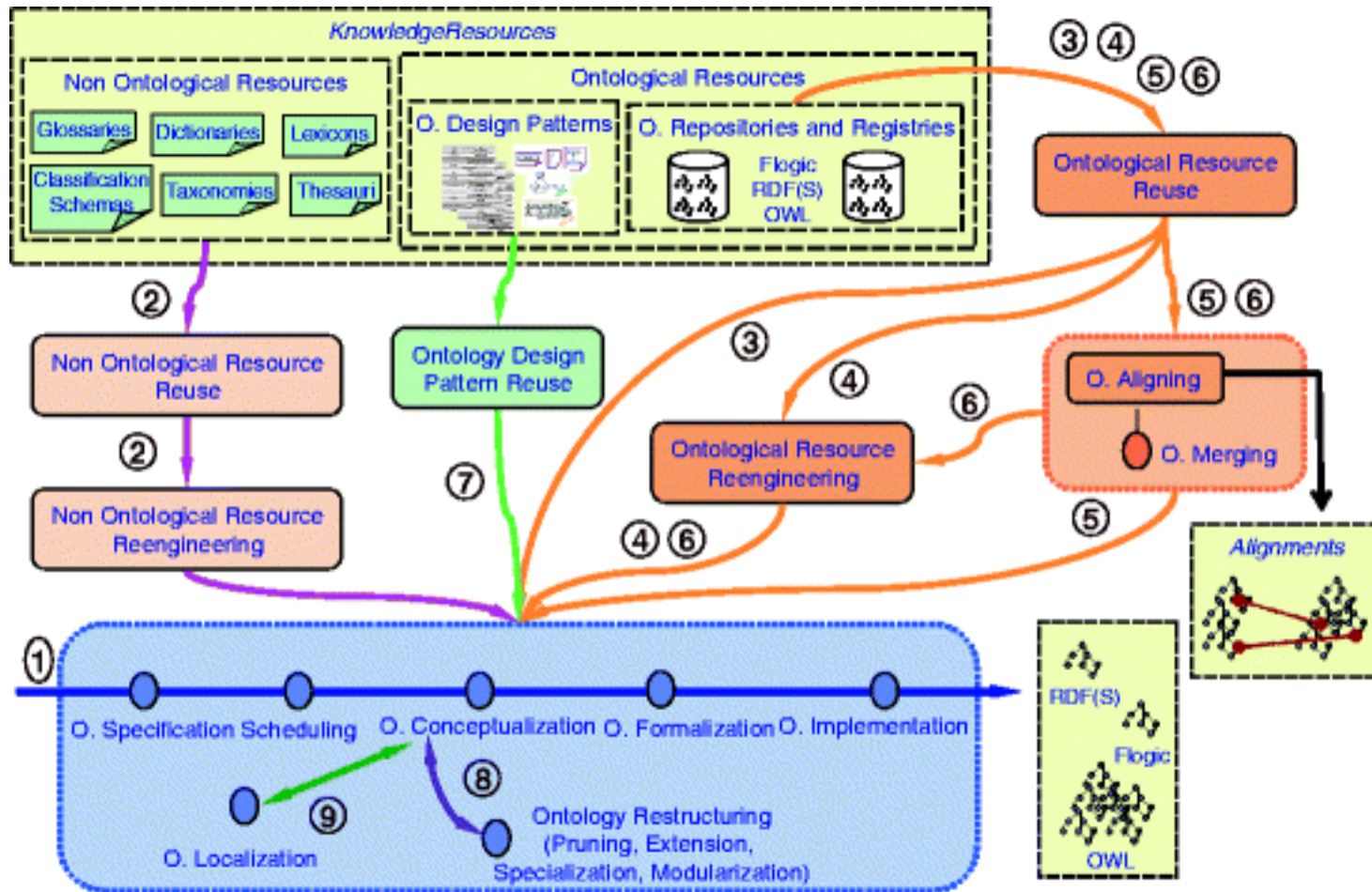
INRAE

Formalisation des connaissances
Novembre 2023 / C Roussey

Ingénierie des ➤ connaissances

Ingénierie des ontologies

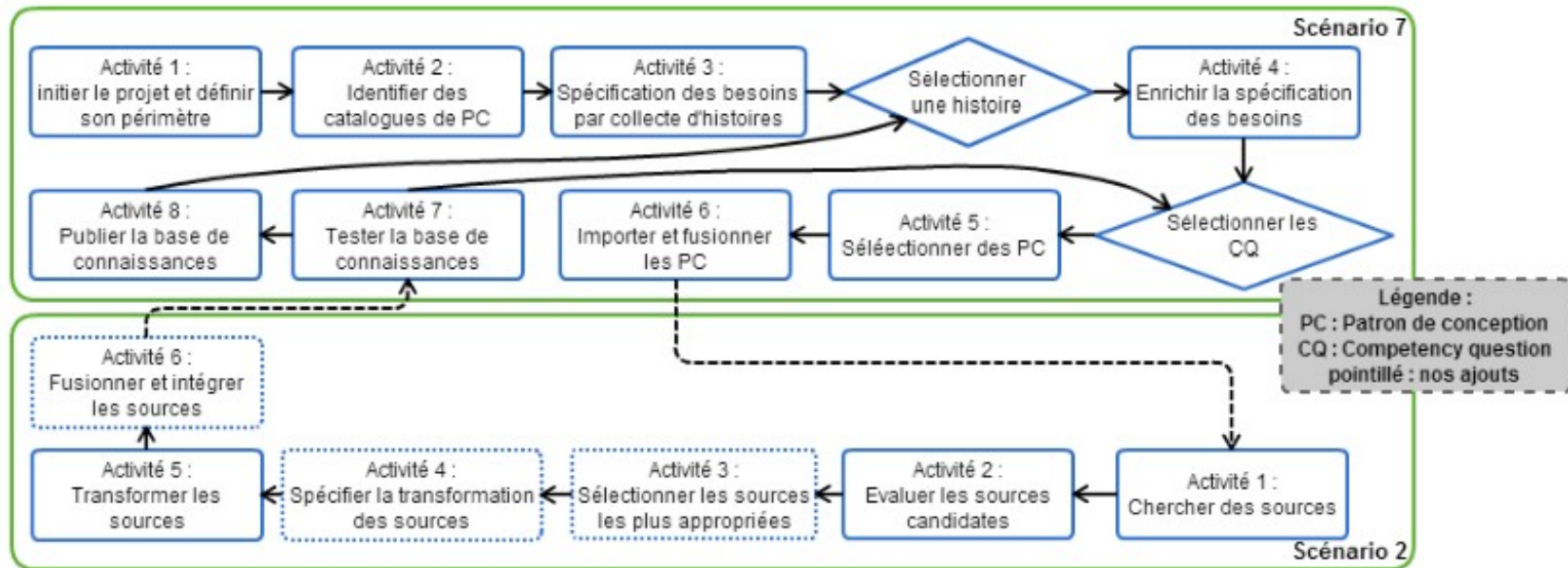
la méthodologie NeOn = un ensemble de méthodes



Crédit Suárez-Figueroa M.C., Gómez-Pérez A., Fernández-López M. (2012) The NeOn Methodology for Ontology Engineering. In: Suárez-Figueroa M., Gómez-Pérez A., Motta E., Gangemi A. (eds) Ontology Engineering in a Networked World. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-24794-1_2

Ingénierie des ontologies

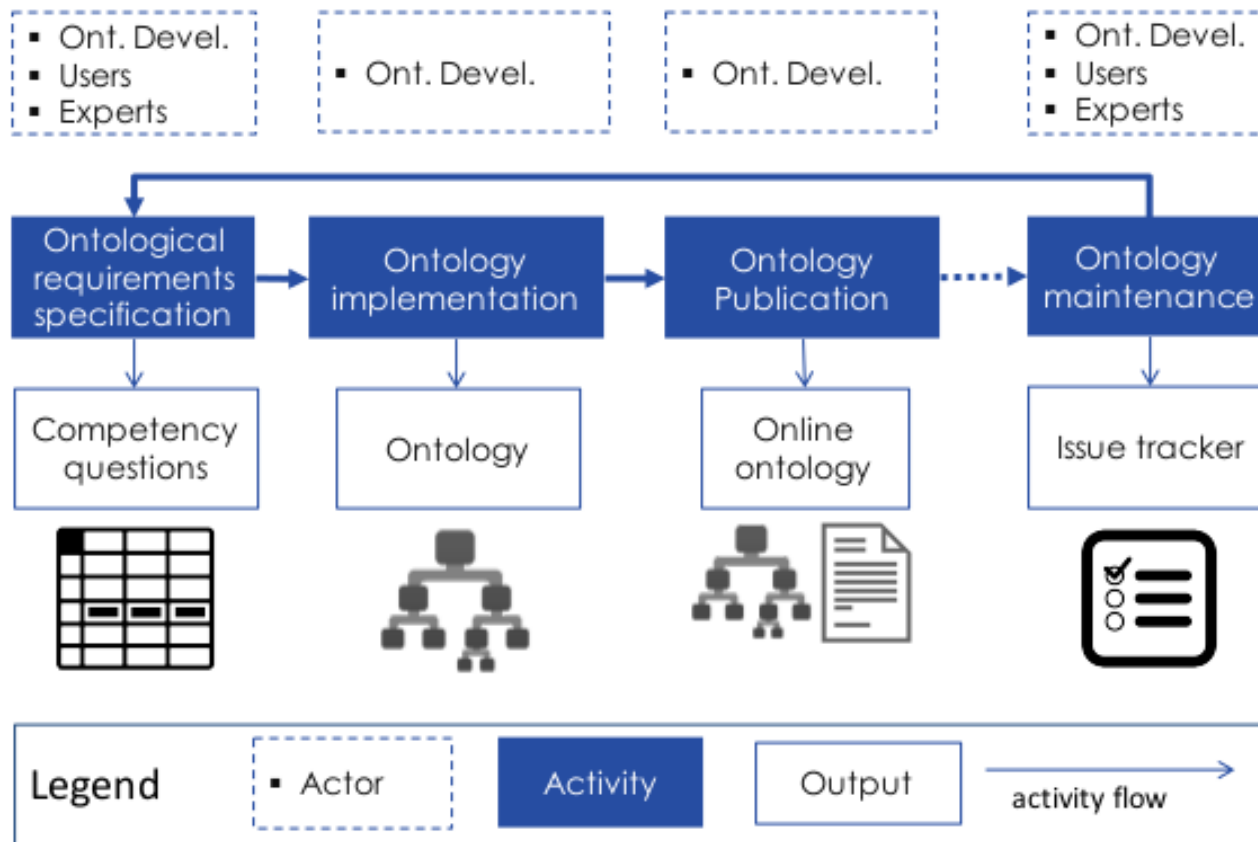
Neon propose plusieurs scénarios que l'on peut combiner



Crédit Fabien AMARGER, Vers un système intelligent de capitalisation de connaissances pour l'agriculture durable : construction d'ontologies agricoles par transformation de sources existantes
Thèse de doctorat en informatique, 2015, Université Toulouse 2

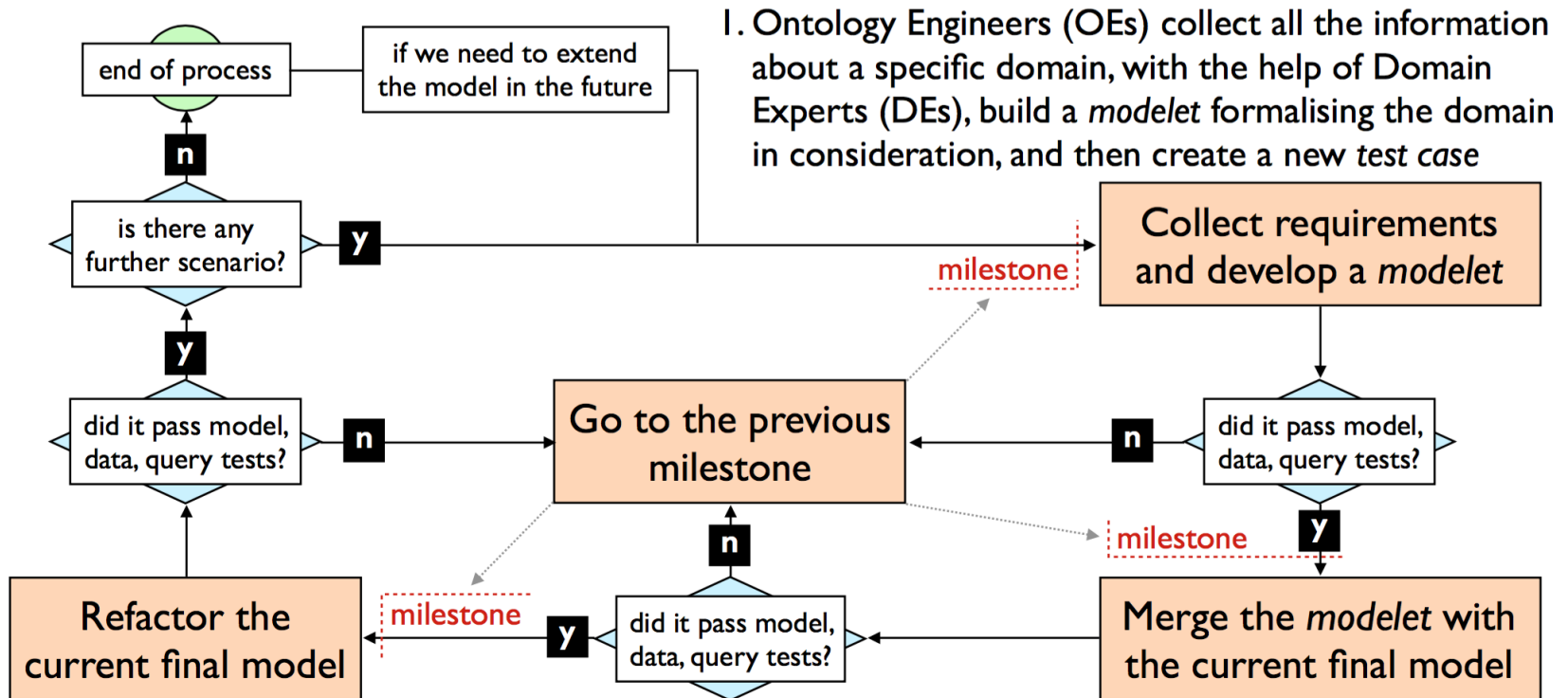
Ingénierie des ontologies

La méthode Linked Open Terms (LOT)



Ingénierie des ontologies

La méthode SAMOD



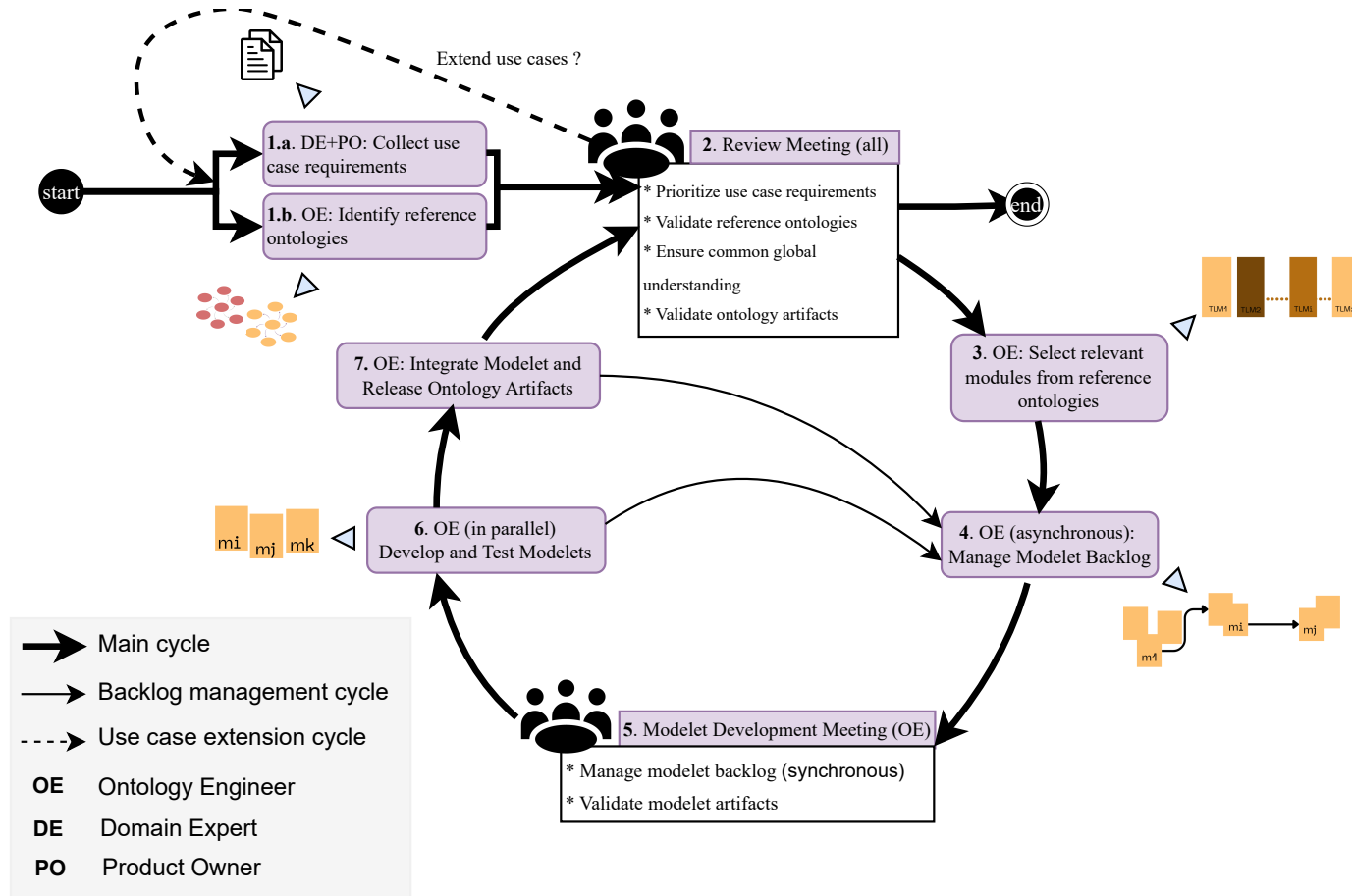
1. Ontology Engineers (OEs) collect all the information about a specific domain, with the help of Domain Experts (DEs), build a *modelet* formalising the domain in consideration, and then create a new test case

3. OEs refactor (important: reuse existing knowledge) the current model, focussing on the last part added in the previous step

2. OEs merge the *modelet* of the new test case with the current model produced by the end of the last process iteration (first iteration: *modelet* becomes current model)

Ingénierie des ontologies

La méthode ACIMOV : une évolution de SAMOD



Crédit F-Z. HANNOU, V. CHARPENAY, M. LEFRANÇOIS, C. ROUSSEY, A. ZIMMERMANN, F. GANDON.
The ACIMOV Methodology: Agile and Continuous Integration for Modular Ontologies and Vocabularies.
In Proceedings of the 2nd Workshop on Modular Knowledge (MK2023), Jul 2023, Sherbrooke, Canada.

Ontology Design Pattern

Design Pattern : une modélisation de petite taille répondant à un problème précis

ODP sont présents :

- dans les ontologies fondationnelles BFO ou DoI Ce
- Dans les ontologies coeur de domaine (core domain) comme Semantic Sensor Network

Un site web dédié ontologydesignpatterns.org et des conférences associées

Ils peuvent prendre différents noms et forme:

- Le Module AgronomicTaxon pour la Thèse de Fabien Amarger
- Le modelet dans la méthode SAMOD



Où trouver des ontologies FAIR ?

Organismes de normalisation

- World Wide Web Consortium (W3C) <https://www.w3.org/>
- European Telecom Standard Institut (ETSI) <https://www.etsi.org/>

Répertoires

- Linked Open Vocabulary (LOV) <https://lov.linkeddata.es/dataset/lov/>
- EMBL-EBI Ontology Lookup Service (OLS) -> biomedical
<https://www.ebi.ac.uk/ols/index>
- Agroportal -> agriculture, environnement, INRAE , ... <http://agroportal.lirmm.fr/>

Communauté de pratique

- Community group CROP ontology <https://www.cropontology.org/>
- Basic Formal Ontology <https://basic-formal-ontology.org/>



Publication de données

➤ capteurs sur le web de données liées

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



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

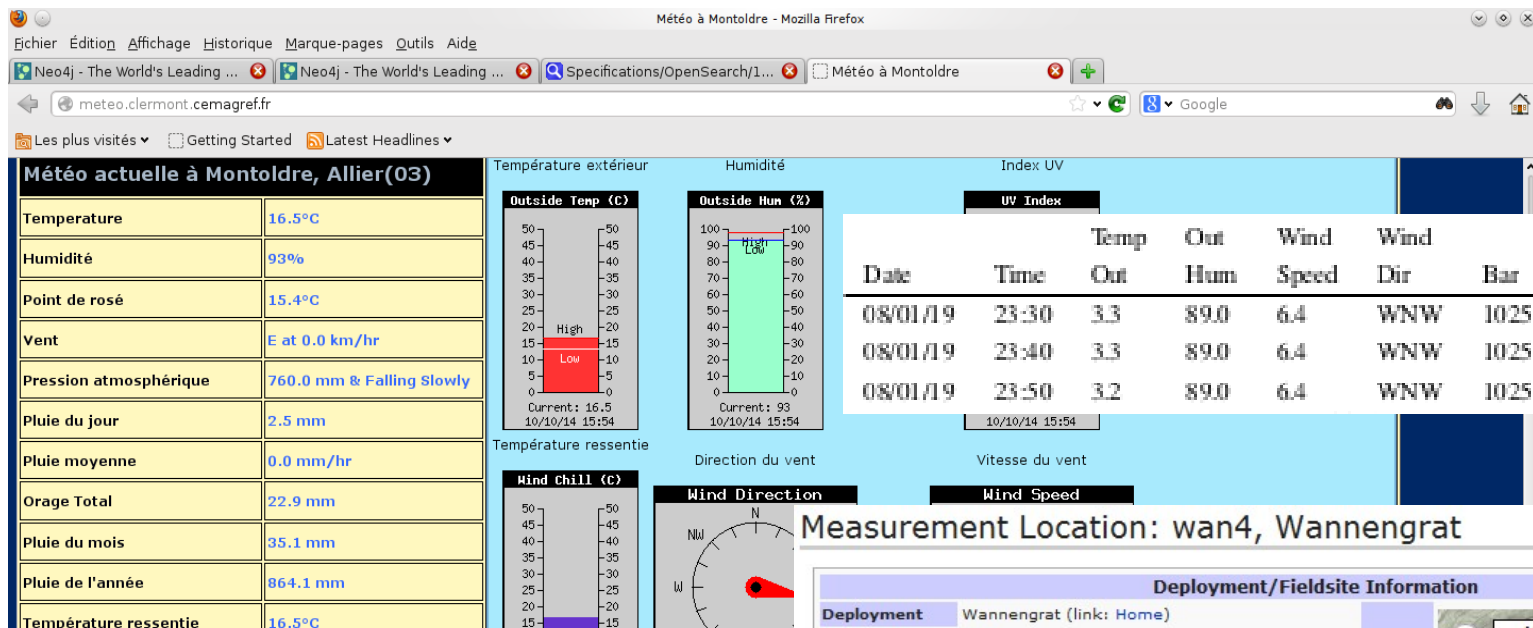
Différents formats / schéma de données / interface d'interrogation

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

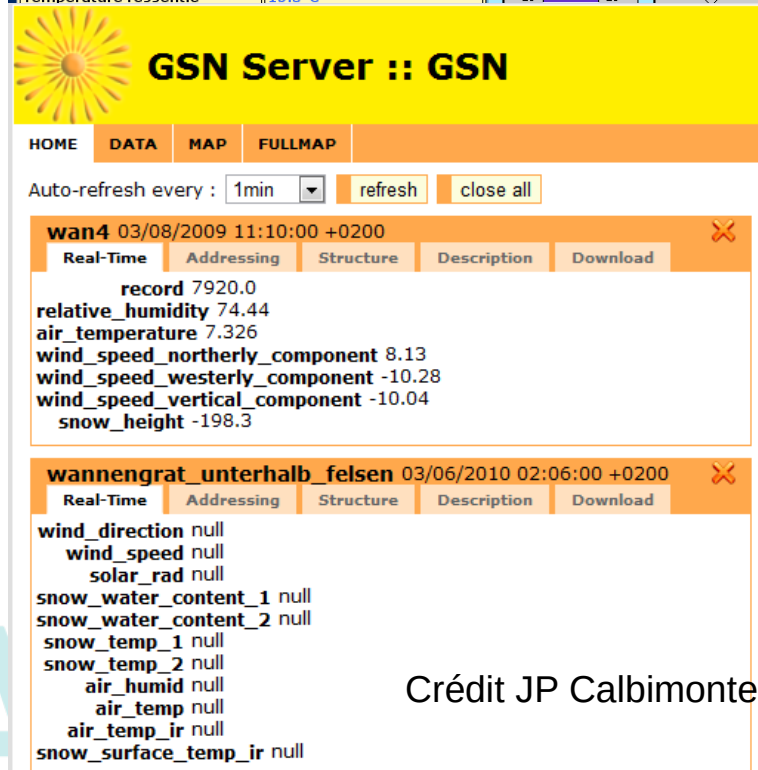
→ Difficulté à interroger de manière unifiée les données issues de différentes stations

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



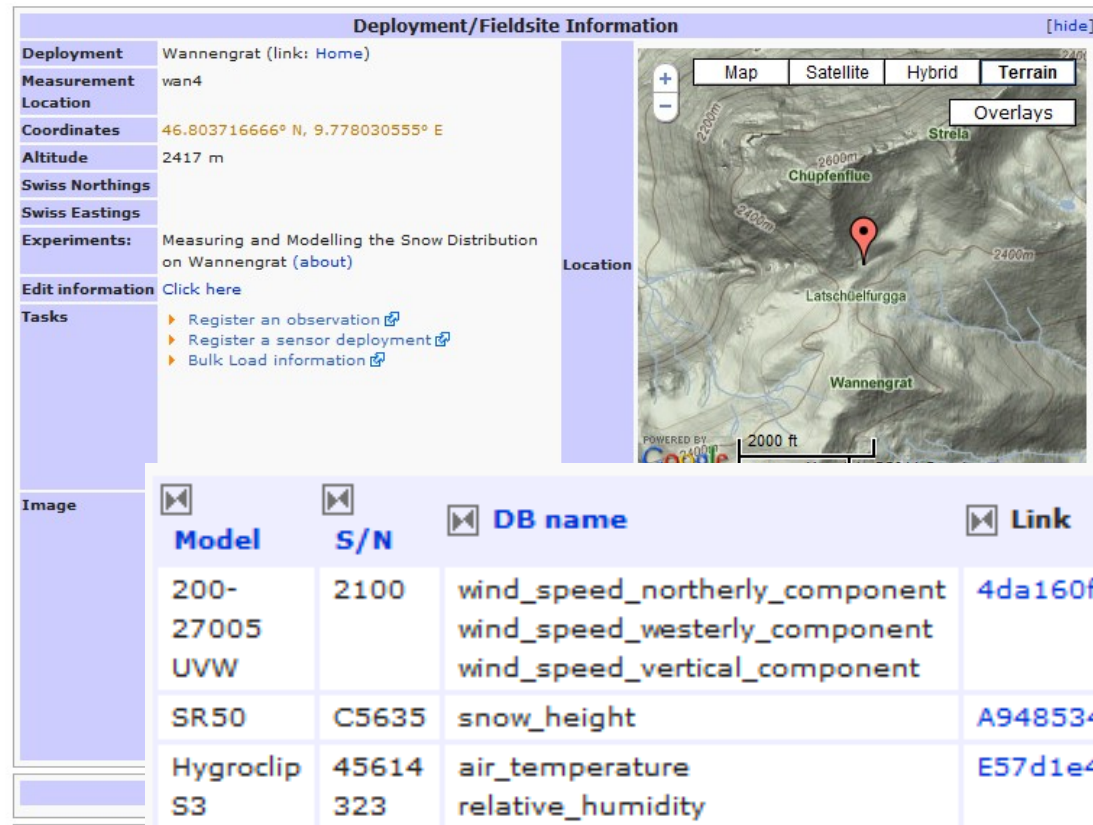


Crédit G André



Crédit JP Calbimonte

Measurement Location: wan4, Wannengrat



Publication de données capteurs

Accès uniformisé aux données et métadonnées des capteurs

« Standardisation » des données:

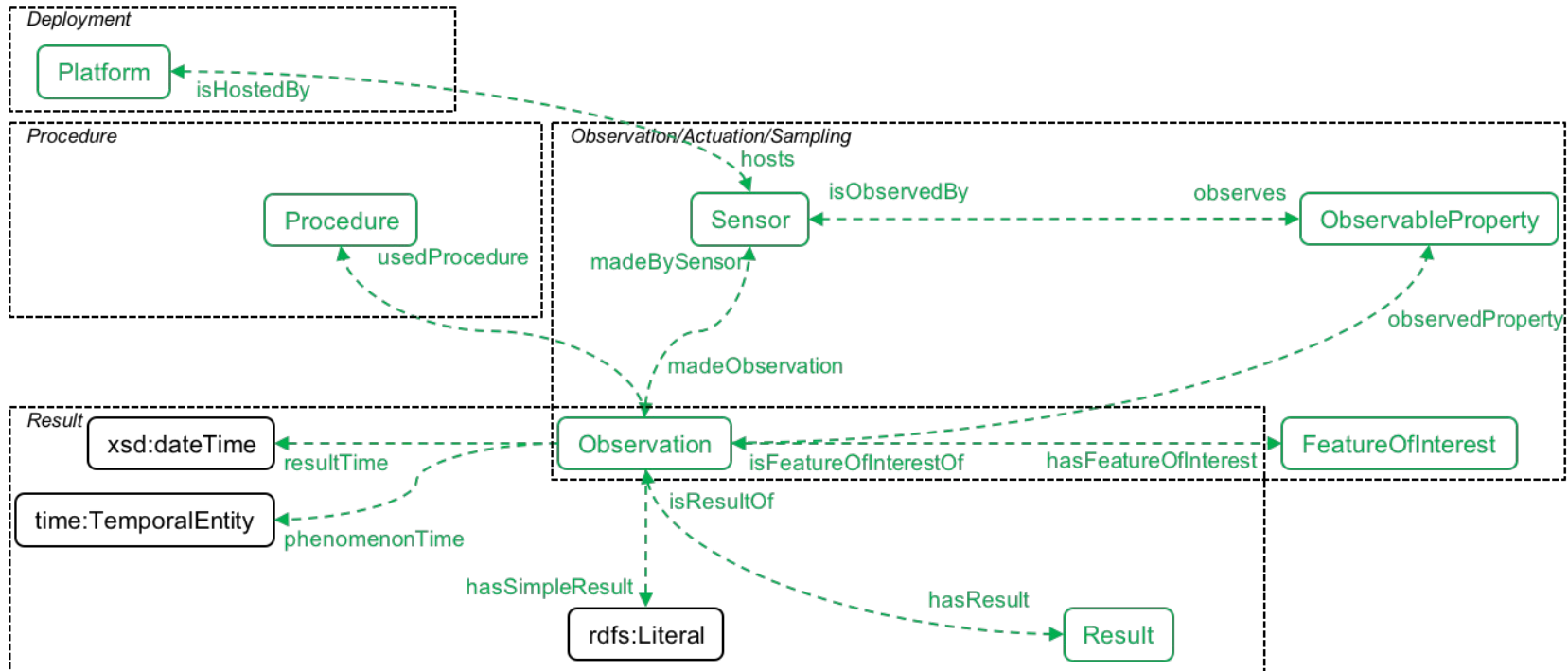
- Réutilisation de schémas (ontologies) existants déjà documentés
- Combinaison de différents schémas pour coller à notre besoin
- Facilité l'intégration de nos données avec d'autres jeux de données
- Facilité la découverte de notre jeu de données (visibilité)

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

Formalisation des connaissances
Novembre 2023 / C Roussey

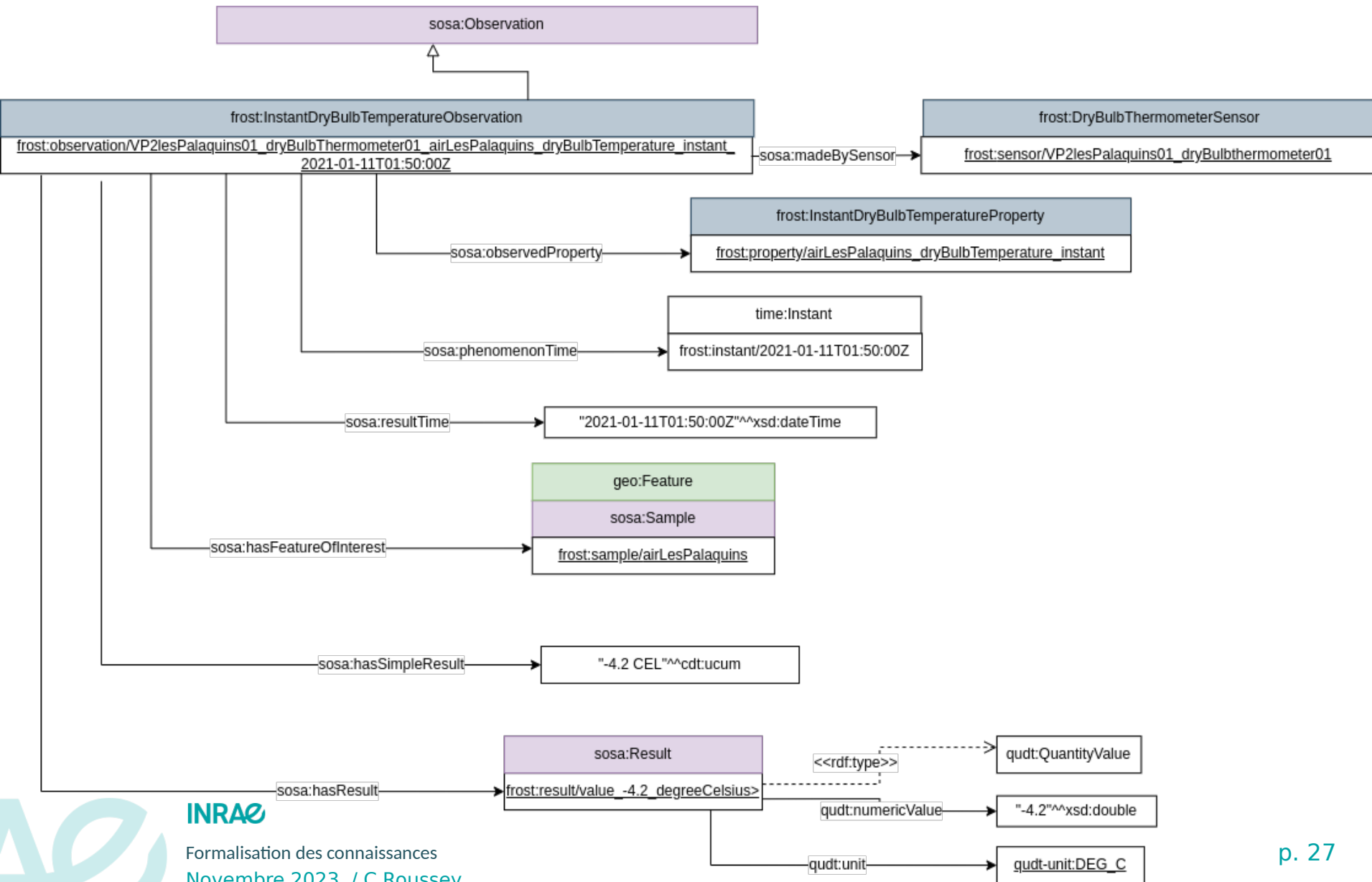
Réseau d'Ontologies

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)

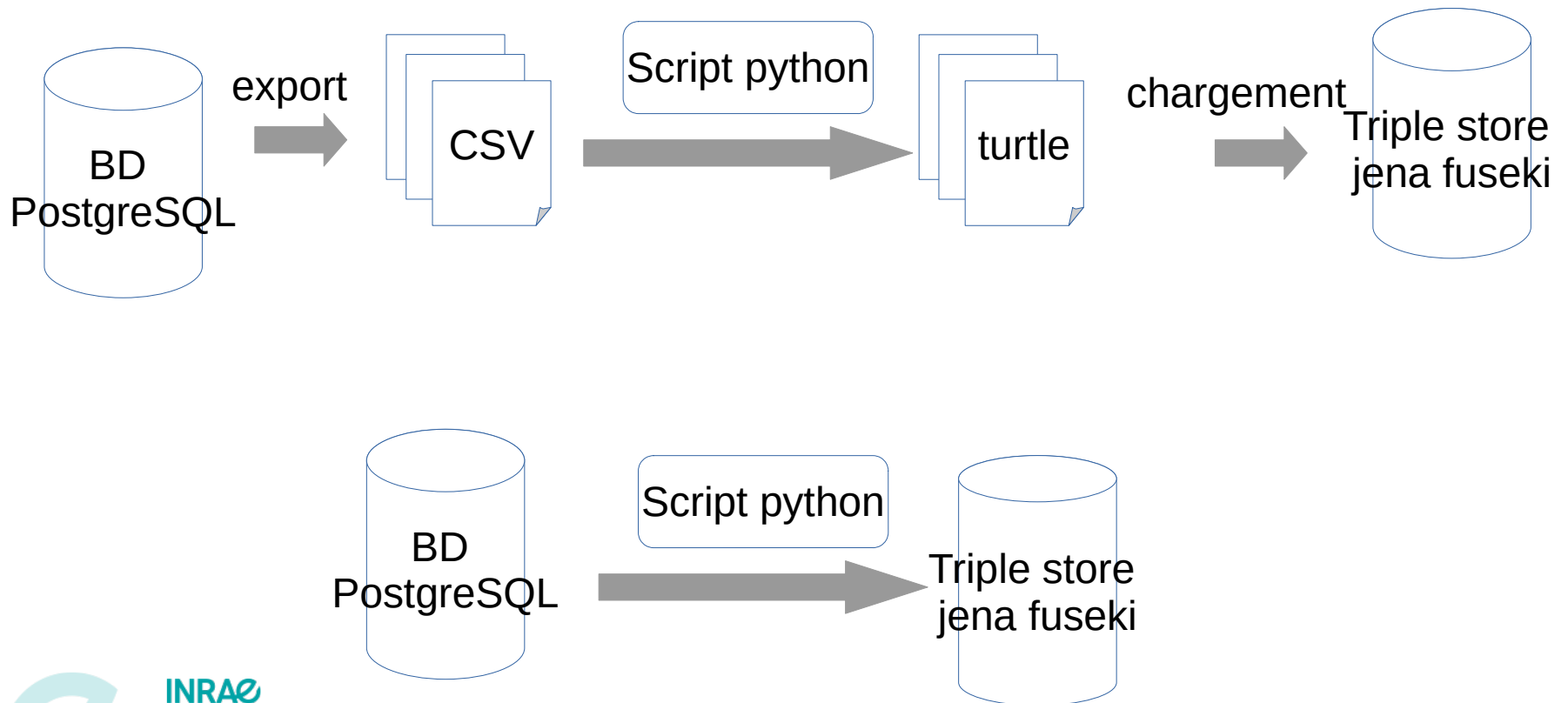


Observation



BD to triple store

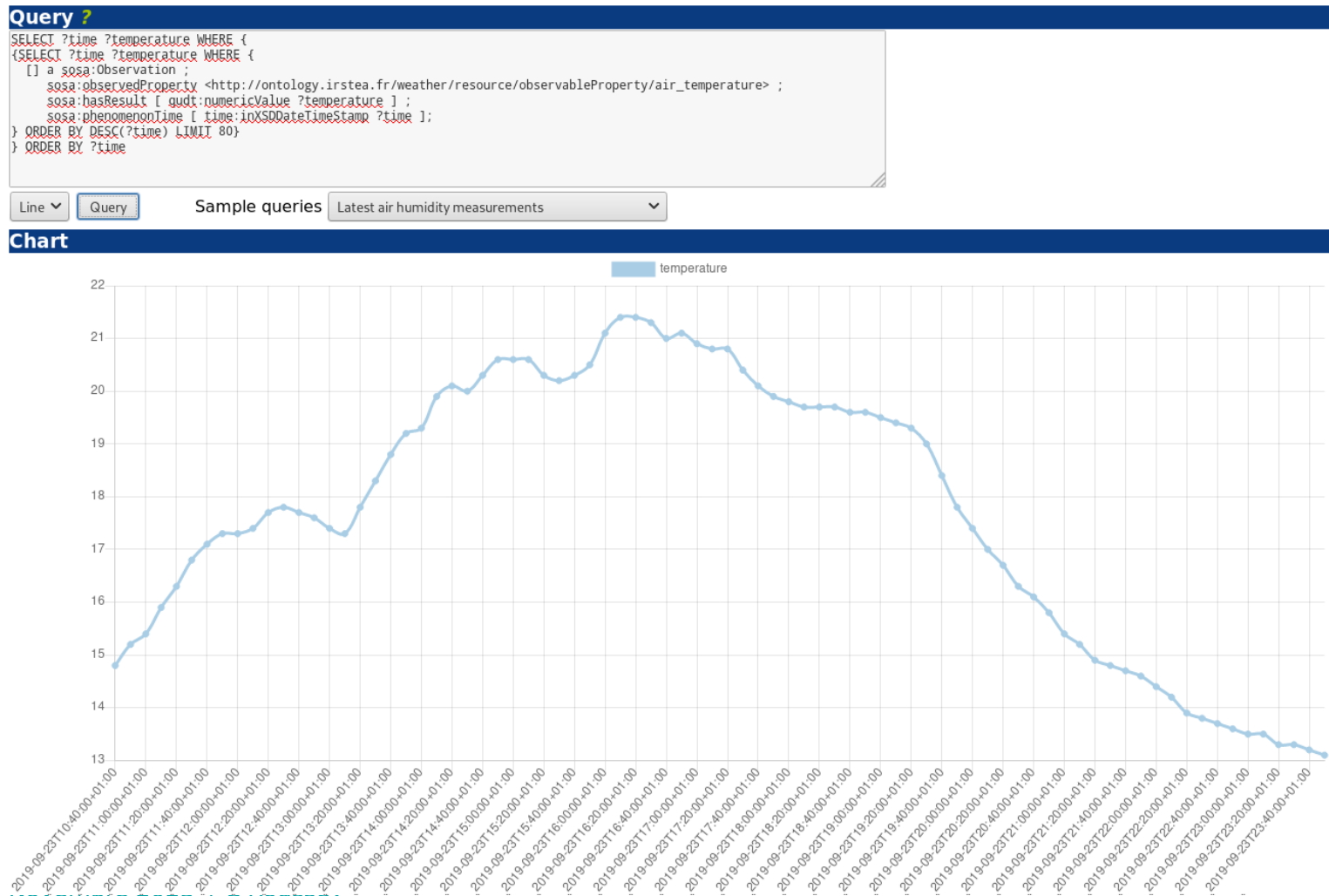
Credit Stephan Bernard stephan.bernard@inrae.fr



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/>



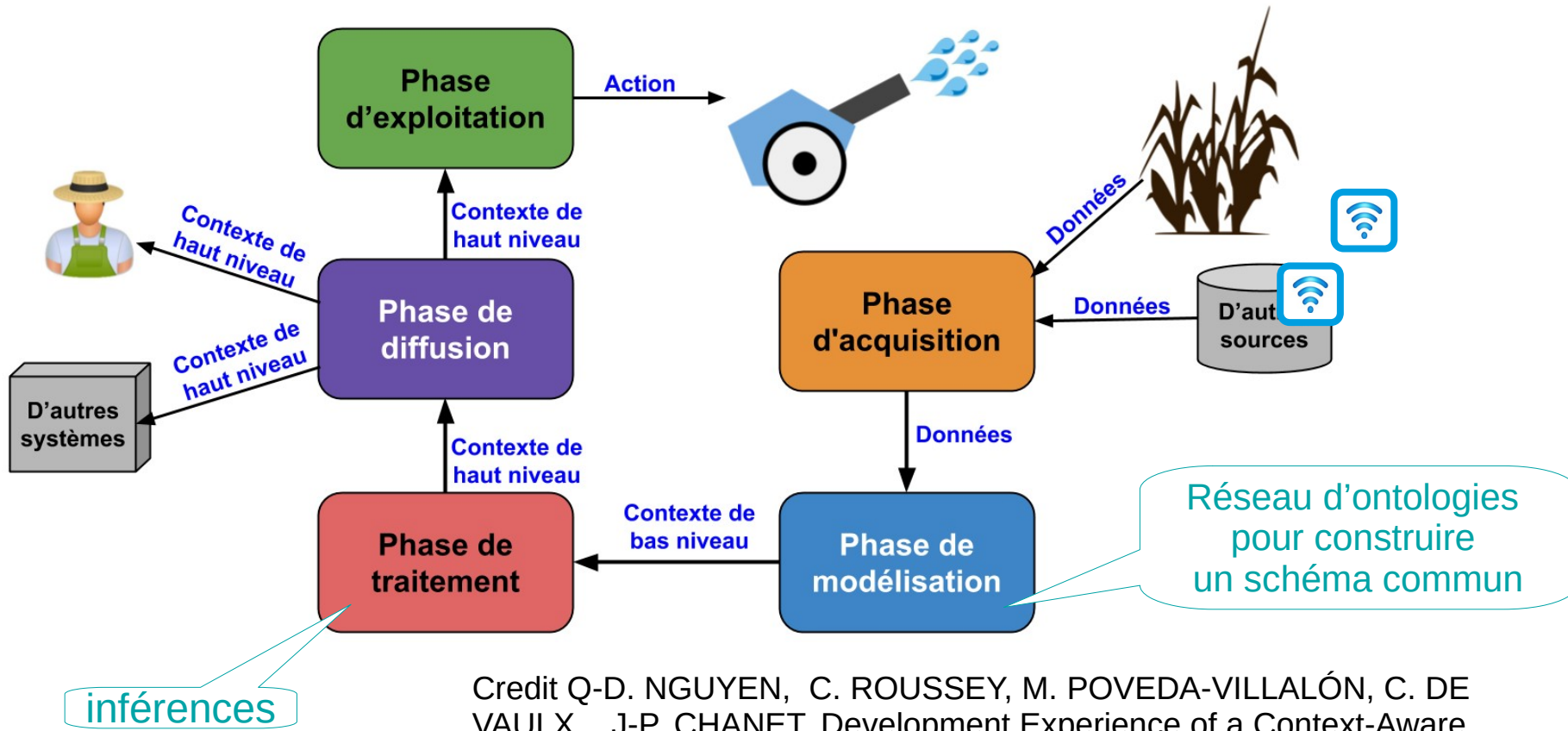


Ontologies pour les ➤ systèmes contextuels adaptatifs

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 → decision → action → observation



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>

Systeme contextuel adaptatif

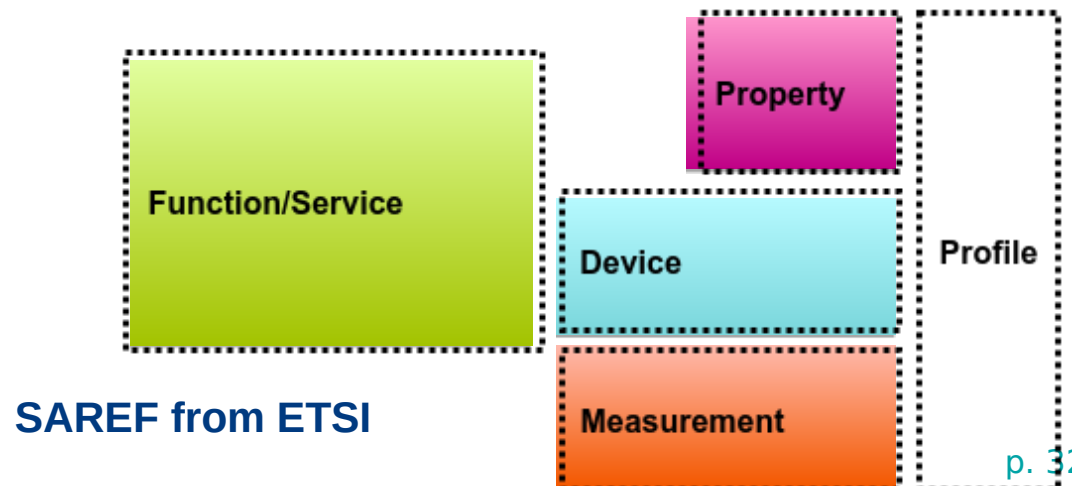
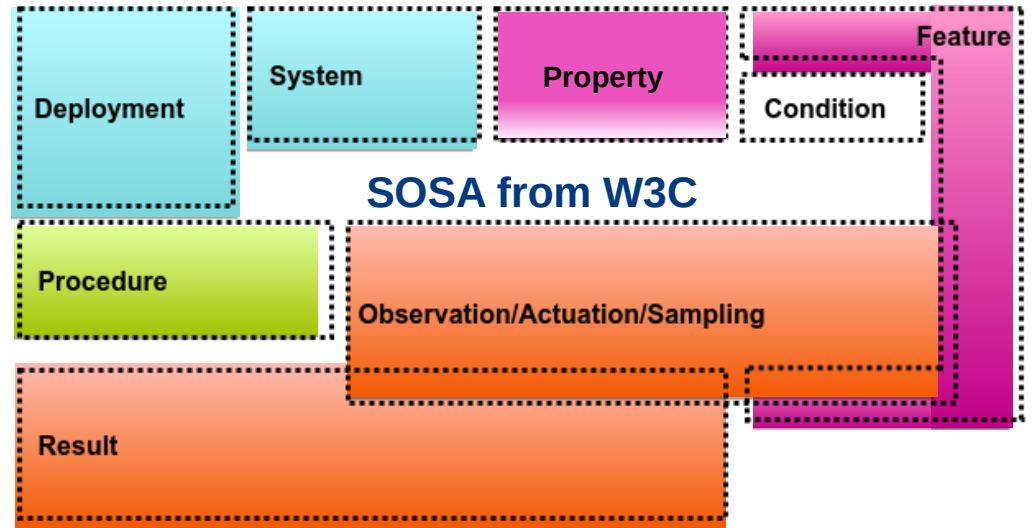
Combiner ontologie d'observation + ontologie de service

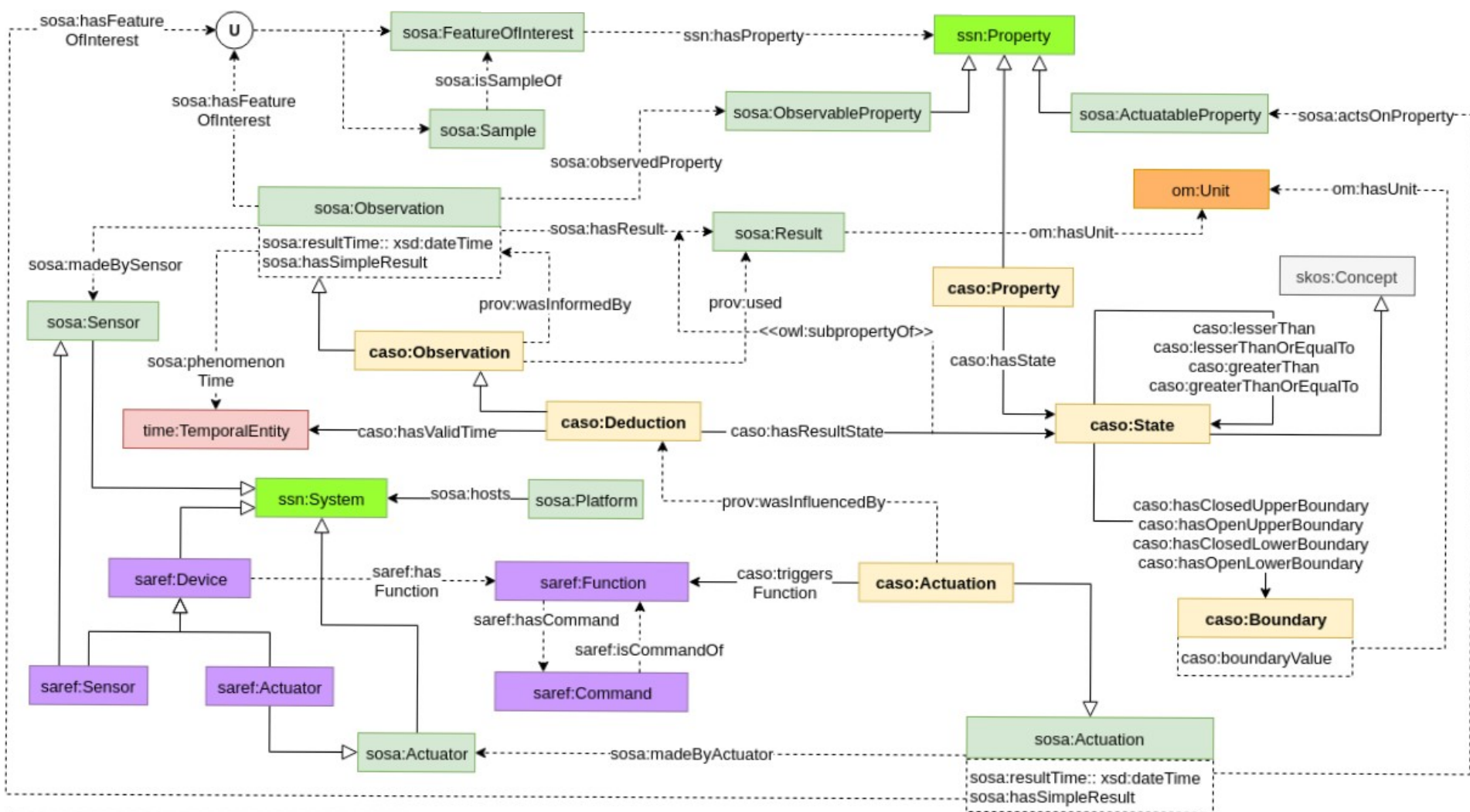
SSN décrit une situation de mesure :
qui, quoi, quand, comment.

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.





Légende

Classe
Attribut dont "domain" est la classe attachée

- sous classe de —>
- propriété d'objet avec "domain" et "range" définis —>
- <<stéréotype>> - - ->
- - - propriété d'objet applicable à la classe attachée - - ->

Classe d'Union (Anonyme)

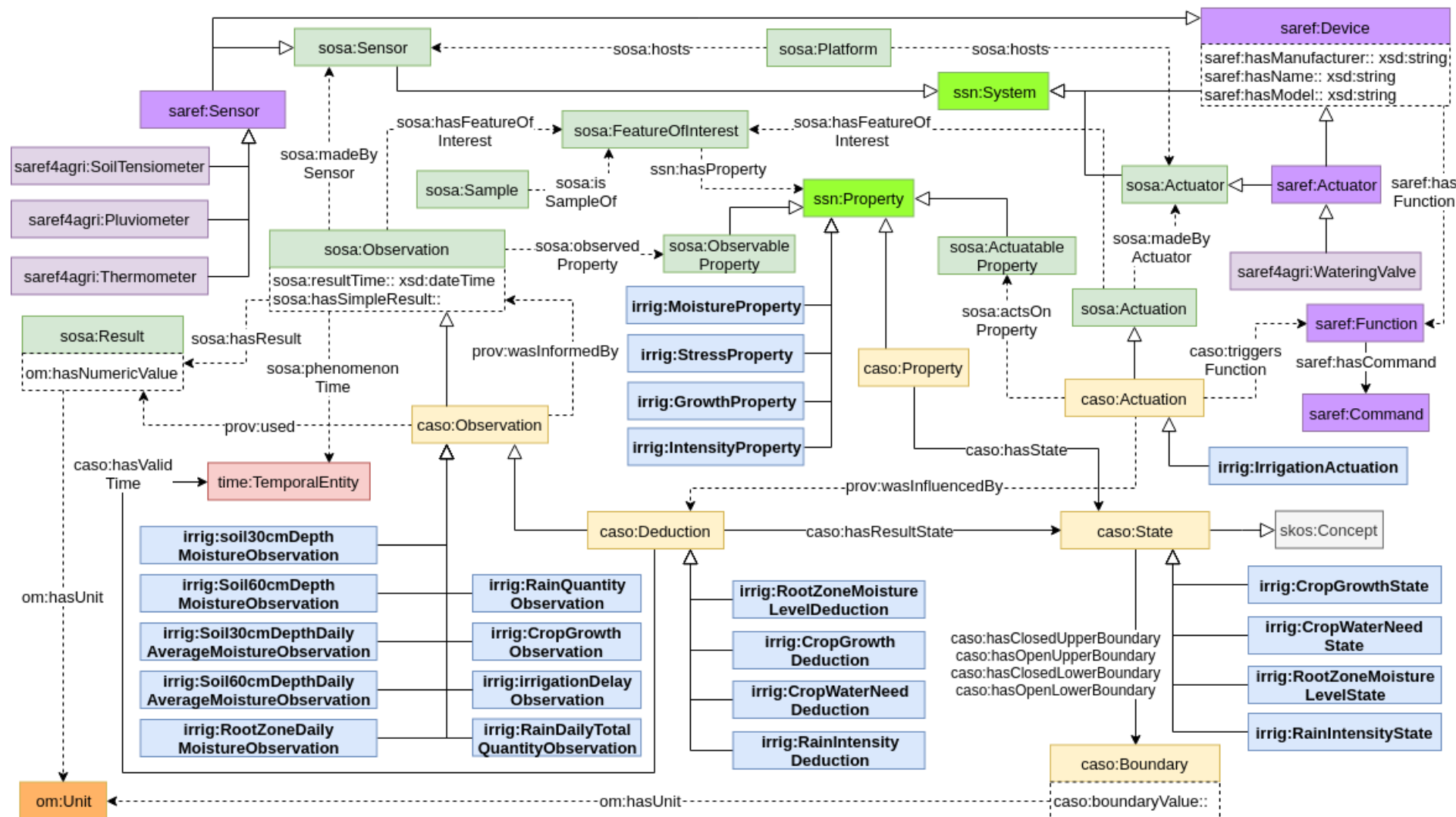


Préfixes

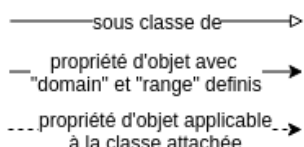
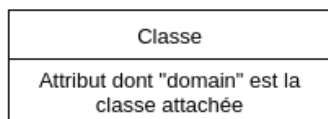
owl: <http://www.w3.org/2002/07/owl#>
sosa: <http://www.w3.org/ns/sosa/>
ssn: <http://www.w3.org/ns/ssn/>
saref: <http://www.w3id.org/saref#>
time: <http://www.w3.org/2006/time#>
om: <http://www.ontology-of-units-of-measure.org/resource/om-2/>
prov: <http://www.w3.org/ns/prov>
skos: <http://www.w3.org/2008/05/skos>

Ontology

caso: <https://w3id.org/def/caso#>



Légende



Préfixes

owl: <https://www.w3.org/2002/07/owl#>
 caso: <https://w3id.org/def/caso#>
 sosa: <http://www.w3.org/ns/sosa/>
 ssn: <http://www.w3.org/ns/ssn/>
 saref: <http://www.w3id.org/saref#>
 saref4agri: <http://www.w3id.org/def/saref4agri#>
 time: <http://www.w3.org/2006/time#>
 om: <http://www.ontology-of-units-of-measure.org/resource/om-2/>
 skos: <http://www.w3.org/2008/05/skos#>

Ontology

irrig: <https://w3id.org/def/irrig#>

Conception: une méthode d'aggrégation pour la pluie

$\text{getRainDailyTotalQuantity}(p, d) = \sum \text{getRainQuantity}(p, t1, t2)$
on the condition that $[t1, t2] \subset [d \text{ 06:00:00}, d+1 \text{ 06:00:00}[$

where

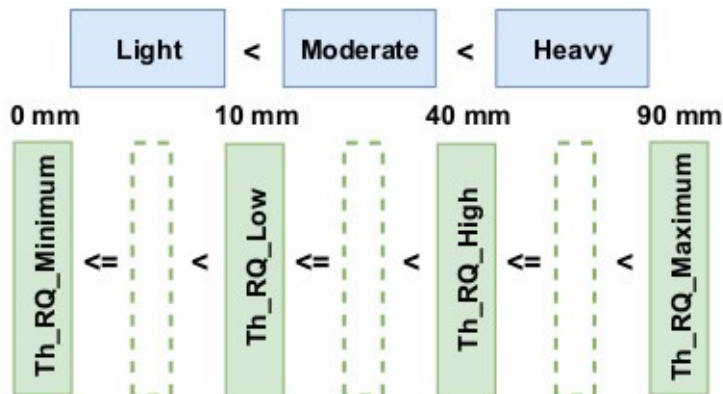
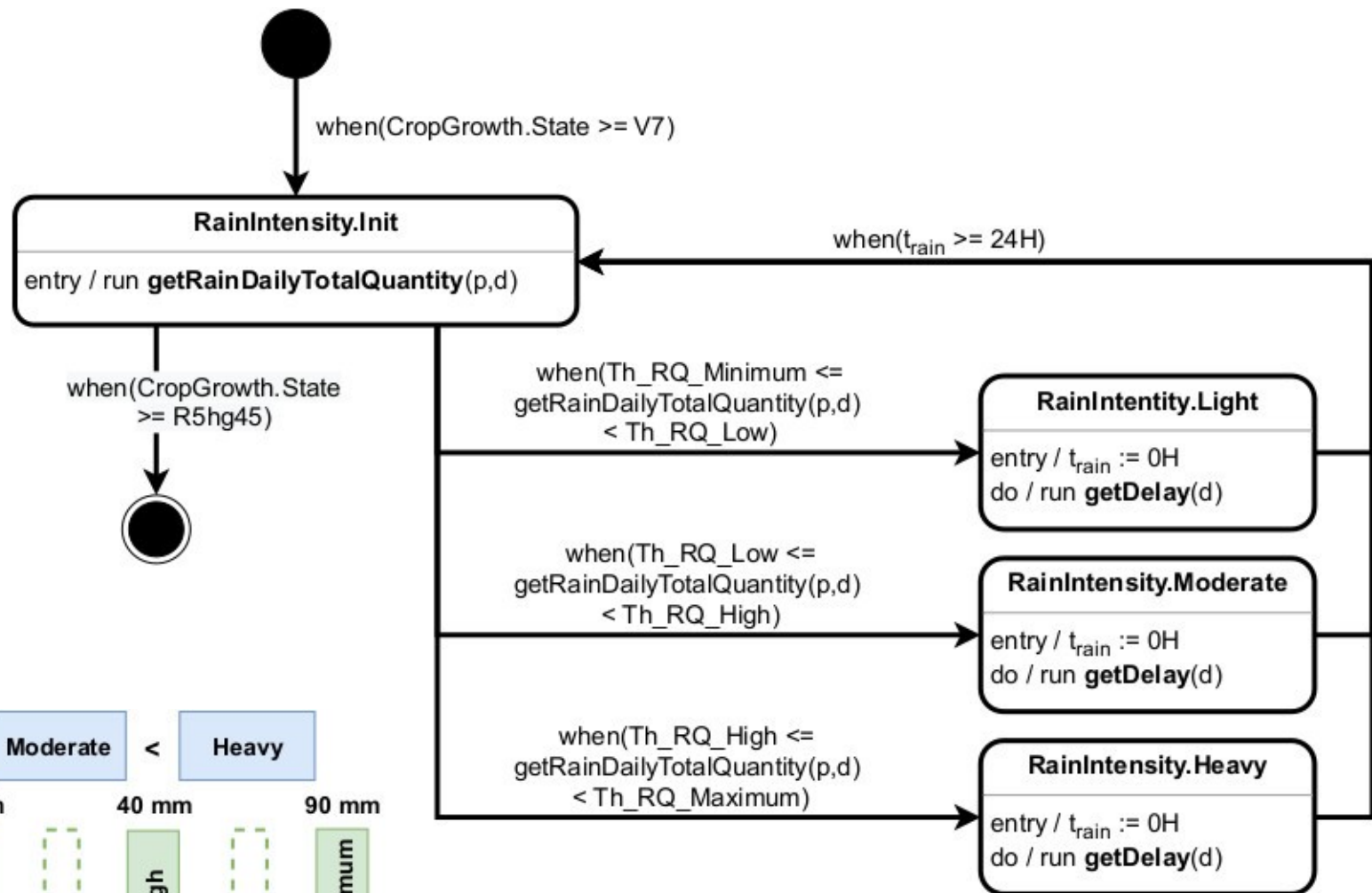
d: a specific day, e.g., 14/08/2013

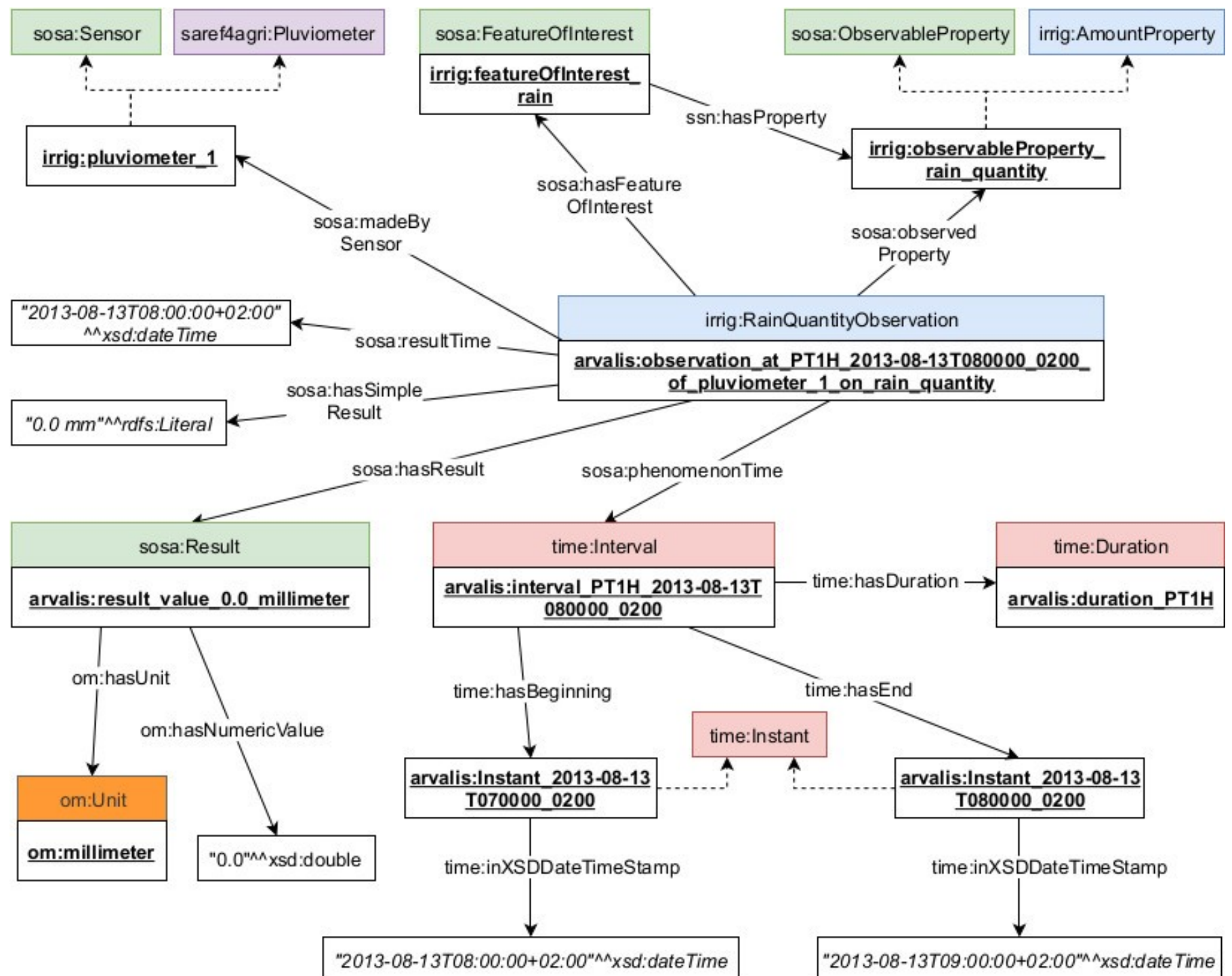
p: a pluviometer

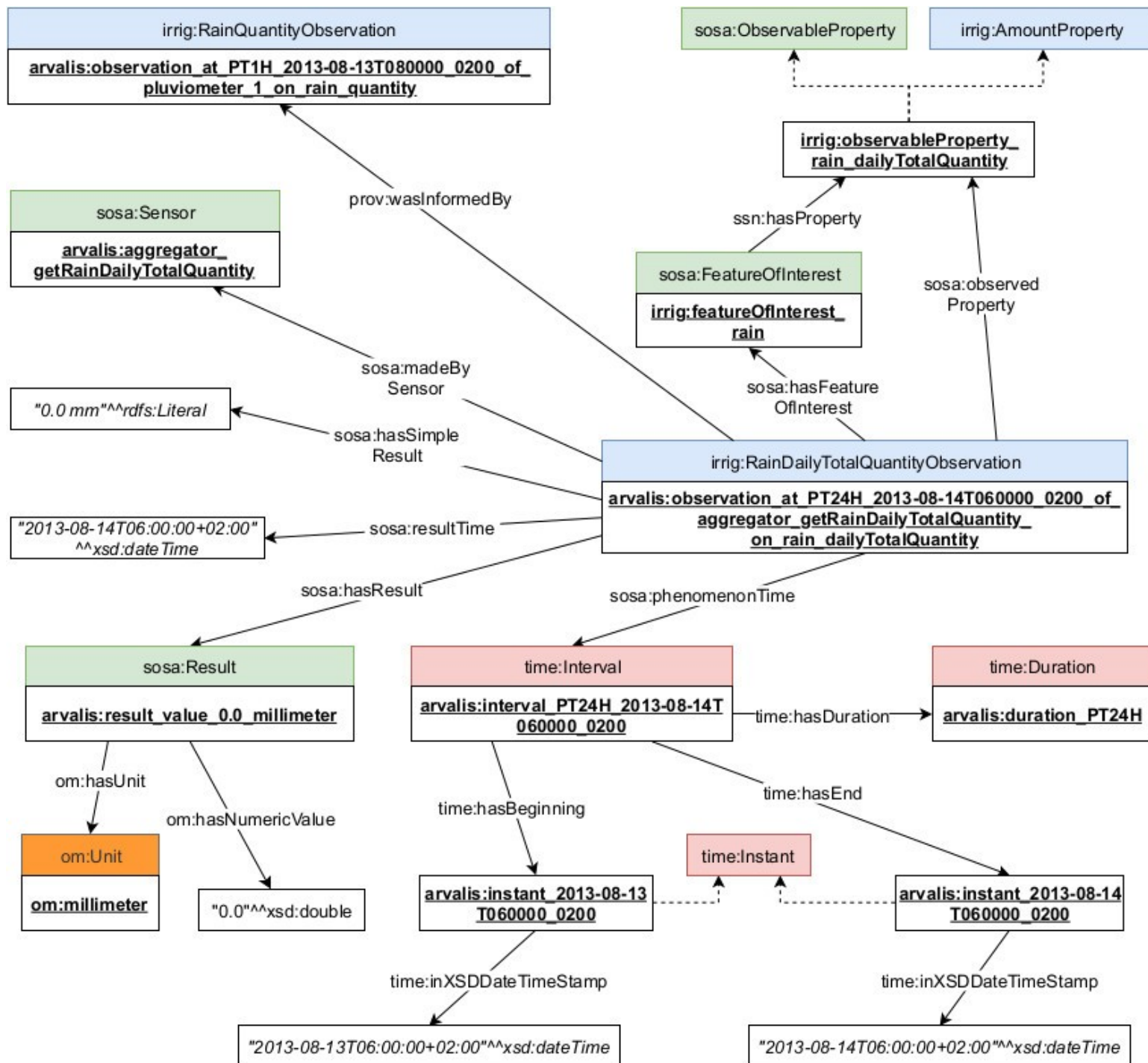
ti : a specific instant of day d, e.g., 08:00:00 on 14/08/2013

$\text{getRainQuantity}(p, t1, t2)$: the rainfall quantity value measured by the pluviometer p during the interval $[t1, t2 [$.

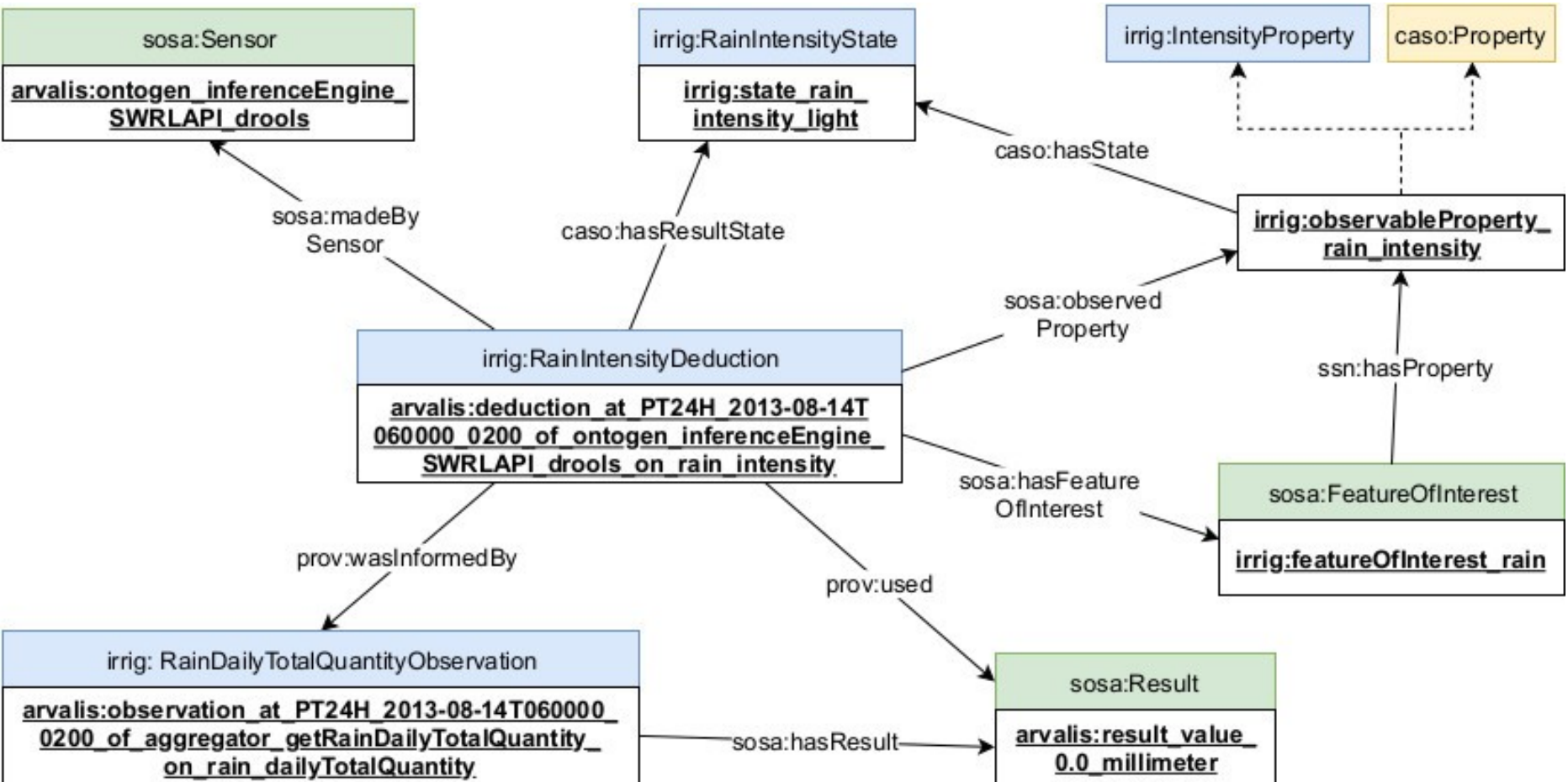
Conception: : diagramme d'état pour déduire les états de la pluie

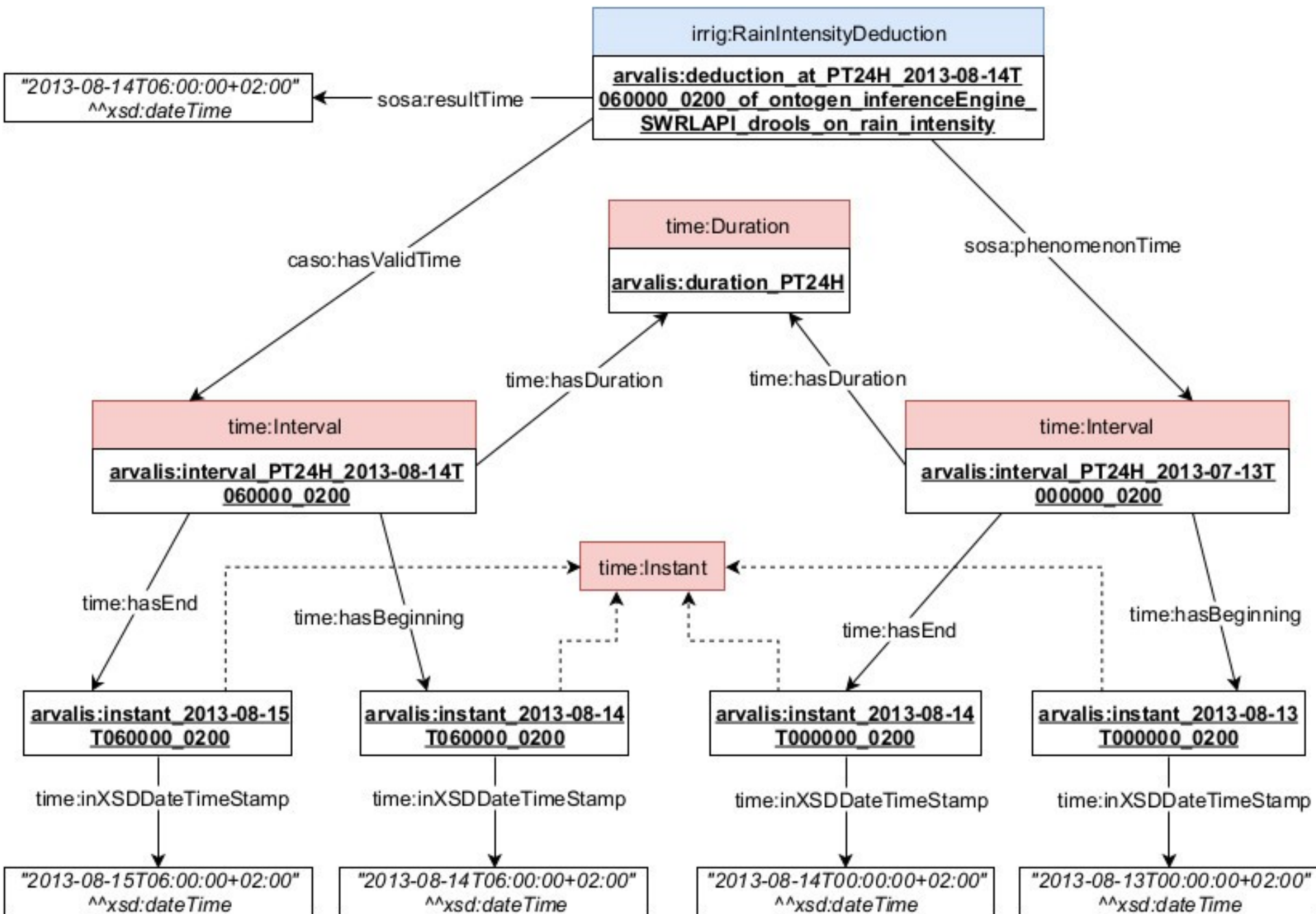






Pluie: Observation et Deduction

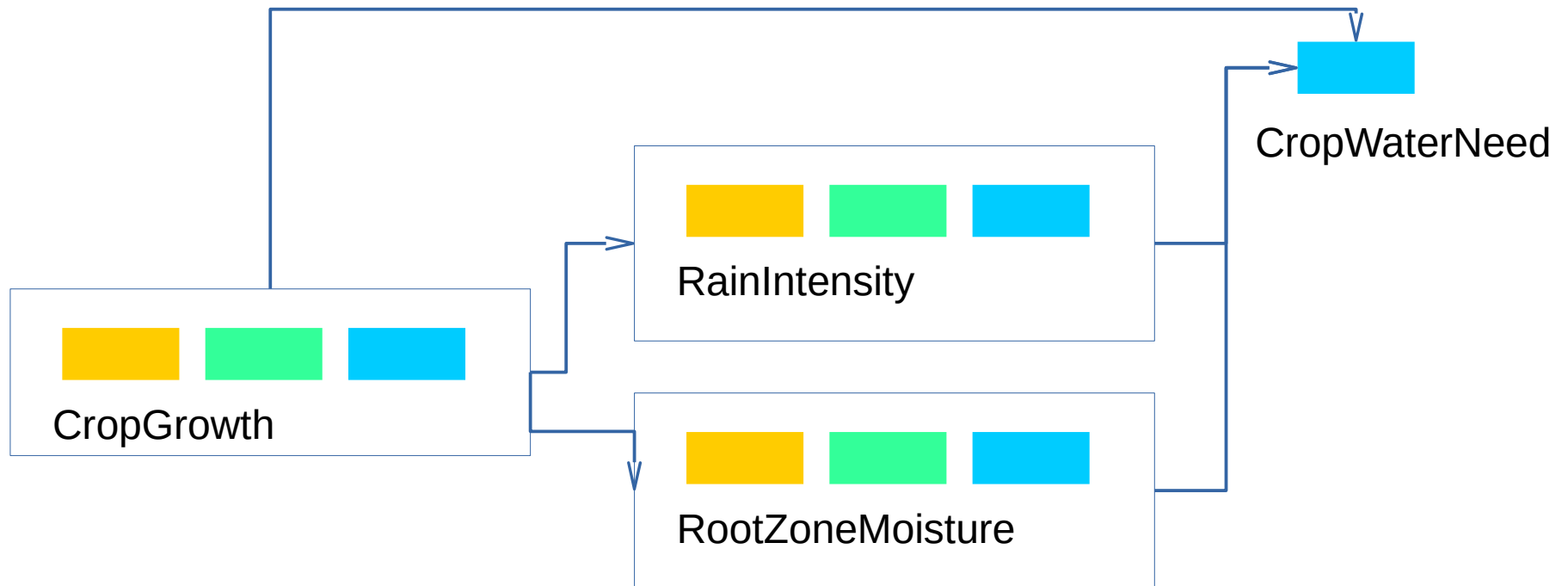




Validation: Une seule règle pour déduire l'état de la pluie

```
irrig:RainIntensityDeduction(?deduction_rain_intensity) ^  
prov:used(?deduction_rain_intensity, ?result_observation_rain_total_quantity) ^  
om:hasNumericalValue(?result_observation_rain_total_quantity, ?  
value_result_observation) ^  
caso:hasState(irrig:observableProperty_rain_intensity, ?rain_intensity_state) ^  
caso:hasOpenUpperBoundary(?rain_intensity_state, ?boundary_upper) ^  
caso:boundaryValue(?boundary_upper, ?value_boundary_upper) ^  
caso:hasClosedLowerBoundary(?rain_intensity_state, ?boundary_lower) ^  
caso:boundaryValue(?boundary_lower, ?value_boundary_lower) ^  
swrlb:lessThan(?value_result_observation, ?value_boundary_upper) ^  
swrlb:greaterThanOrEqual(?value_result_observation, ?value_boundary_lower)  
->  
caso:hasResultState(?deduction_rain_intensity, ?rain_intensity_state)
```

Les chaînes de traitements



Projet D2KAB



> Annotation sémantique de bulletins d'alertes agricoles



D2KAB



Le projet ANR D2KAB 2019-2023 « Data to Knowledge in Agronomy and Biodiversity » (www.d2kab.org) a pour objectif :

- d'améliorer la visibilité et l'échange de jeux de données et de référentiels par le biais de l'Agroportal
- → rendre ces données FAIR par le biais des technologies web sémantique : ontologie, SKOS, graphe de connaissances
- Exploiter ces données FAIR dans des applications dédiées

5 scénarios : alimentaire, agronomie, agriculture, écosystème, biodiversité

Scénario agricole : Navigateur augmenté des Bulletins Agricole

Ressources sémantiques : thésaurus des cultures, échelles phénologiques, variétés des cultures, régions françaises, maladie des plantes et bioagresseurs

Annotations des bulletins pour faciliter la recherche et les analyses

Bulletins de Santé du Végétal (BSV)



BSV n° 18 - 30 juillet 2019

A RETENIR CETTE SEMAINE

Mildiou : Situation bien maîtrisée. Grappes saines. Fin de la protection _____ p 2

Oïdium : Fréquence de parcelles touchées en légère progression _____ p 2

Tordeuses : Vol terminé. Pression G2 très faible. Fin du risque _____ p 3

STADES PHENOLOGIQUES



Les grappes sont maintenant bien fermées, voire compactes. Quelques baies vérées sont signalées au vignoble depuis la fin de la semaine dernière, mais elles restent rares pour l'instant. En raison de l'épisode caniculaire de la semaine dernière, durant lequel des records absolus de températures ont été enregistrés (42.9 °C), de l'échaudage est fréquemment signalé au vignoble. Les dégâts sont en cours d'évaluation.

Le stade 33 « fermeture de la grappe » est observé depuis mi-juillet. Le stade 35 (une grappe sur deux avec au moins une baie vérée) pourrait donc être atteint la semaine prochaine.

Chardonnay, Pinot noir, Meunier : stade 33 « fermeture de la grappe ». La phénologie de la vigne reste proche de la moyenne décennale.

Les premiers prélèvements du réseau matu auront lieu le 19 août.

Premières baies vérées

Les BSV constituent une archive d'observations des parcelles agricoles françaises et décrivent :

un état sanitaire des cultures

stades de développement

observations des ravageurs et maladies

présence de symptômes

une évaluation du risque phytosanitaire, en fonction des périodes de sensibilité des cultures et des seuils de nuisibilité des ravageurs et maladies

des messages réglementaires

INRAE

Formalisation des connaissances
Novembre 2023 / C Roussey

Transformation des pdf en html



BSV n° 18 - 30 juillet 2019

A RETENIR CETTE SEMAINE

Mildiou : Situation bien maîtrisée. Grappes saines. Fin de la protection p 2

Oïdium : Fréquence de parcelles touchées en légère progression p 2

Tordeuses : Vol terminé. Pression G2 très faible. Fin du risque p 3

STADES PHENOLOGIQUES



Premières baies vérées

Les grappes sont maintenant bien fermées, voire compactes.

Quelques baies vérées sont signalées au vignoble depuis la fin de la semaine dernière, mais elles restent rares pour l'instant.

En raison de l'épisode caniculaire de la semaine dernière, durant lequel des records absolus de températures ont été enregistrés (42.9 °C), de l'échaudage est fréquemment signalé au vignoble. Les dégâts sont en cours d'évaluation.

Le stade 33 « fermeture de la grappe » est observé depuis mi-juillet. Le stade 35 (une grappe sur deux avec au moins une baie vérée) pourrait donc être atteint la semaine prochaine.

Chardonnay, Pinot noir, Meunier : stade 33 « fermeture de la grappe ». La phénologie de la vigne reste proche de la moyenne décennale.

Les premiers prélèvements du réseau matu auront lieu le 19 août.

BSV n° 18 - 30 juillet 2019

A RETENIR CETTE SEMAINE

Mildiou : Situation bien maîtrisée. Grappes saines. Fin de la protection p 2

Oïdium : Fréquence de parcelles touchées en légère progression p 2

Tordeuses : Vol terminé. Pression G2 très faible. Fin du risque p 3

STADES PHENOLOGIQUES

FRUIT

Les grappes sont maintenant bien fermées, voire compactes. Quelques baies vérées sont signalées au vignoble depuis la fin de la semaine dernière, mais elles restent rares pour l'instant. En raison de l'épisode caniculaire de la semaine dernière, durant lequel des records absolus de températures ont été enregistrés (42.9 °C), de l'échaudage est fréquemment signalé au vignoble. Les dégâts sont en cours d'évaluation.

Le stade 33 " fermeture de la grappe " est observé depuis mi-juillet. Le stade 35 (une grappe sur deux avec au moins une baie vérée) pourrait donc être atteint la semaine prochaine.

Chardonnay, Pinot noir, Meunier : stade 33 " fermeture de la grappe ". La phénologie de la vigne reste proche de la moyenne décennale.

Les premiers prélèvements du réseau matu auront lieu le 19 août.

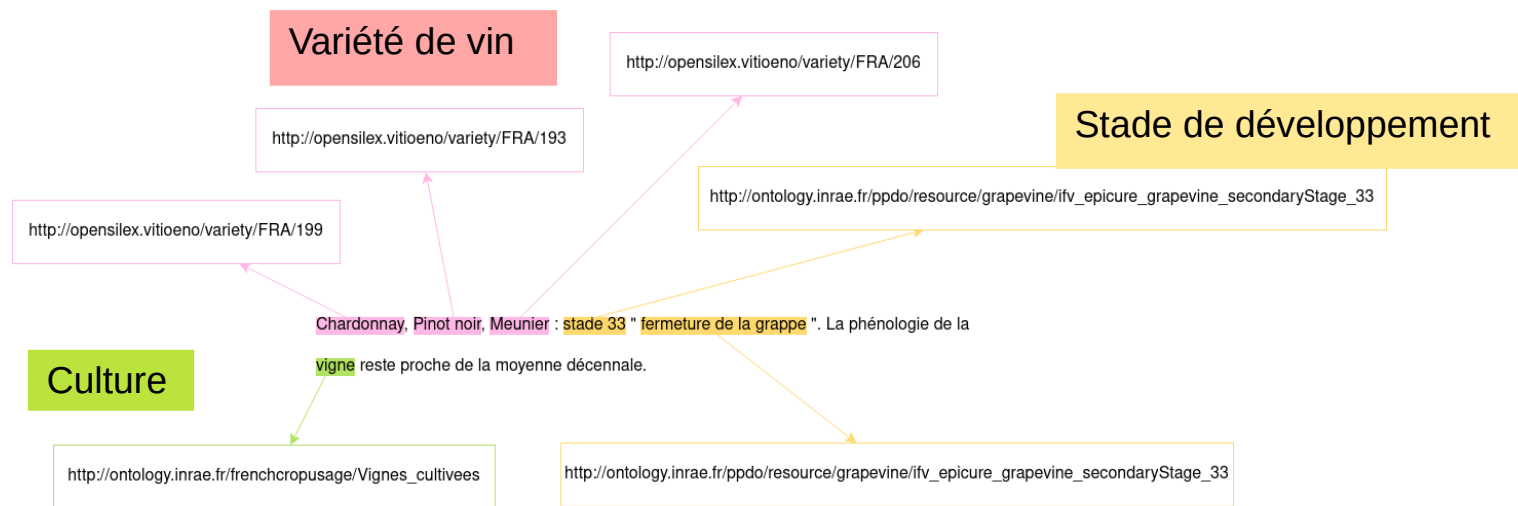
*Premières baies
vérées*

*Le réseau Vigne
AOC Champagne
compte 143
parcelles
observées cette
semaine.*

BSV n° 18 - 1 / 4

Exemple d'Annotation

Observation d'un stade de développement atteint par une variété de vin, dans le bulletin de la région Champagne publié en juillet 2019



Spécification des besoins

Besoins des experts

- Quelles sont les noms de culture mentionnées dans les bulletins de la région Pays de la Loire ? (liste des cultures principales de la région)
- Quelles sont les maladies les plus fréquemment mentionnées dans les bulletins de viticulture en 2019 ? (liste des maladies observées sur une culture à une période connue)
- Pour chaque région publiant un bulletin en viticulture, quelle est la période pendant laquelle le stade Floraison est observé ? (évolution de la culture en fonction des changements climatiques)
- Quels sont les noms scientifiques des bioagresseurs utilisés dans les BSV qui ne sont pas les noms de référence des taxons connus actuellement ? (nom d'usage des agriculteurs ou des agronomes différent des connaissances actuelles de la taxonomie des plantes)

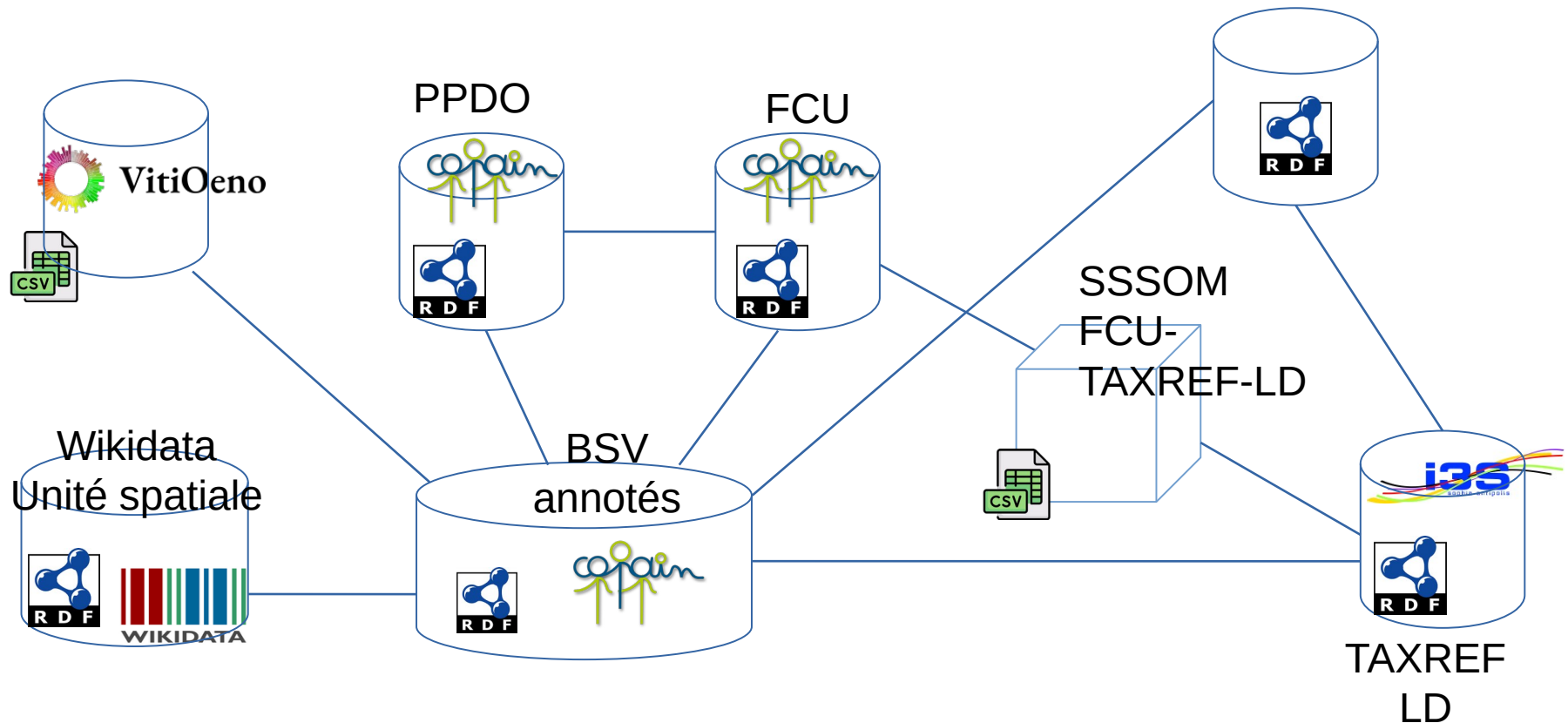
Besoins d'évaluation

- Quelles sont les paires d'annotations automatiques et humaines qui couvrent la même mention mais sont normalisées différemment ? (Quelles sont les annotations automatiques incorrectes?)
- Quels sont les organismes nuisibles qui ont été détectés dans l'annotation manuelle mais pas l'annotation automatique ? (quels sont les organismes nuisibles non défini dans la source de référence?)



Graphes distribués

Interroger des graphes de connaissances distribués sur le web ET liés les uns aux autres
NCBITaxon



Thésaurus usage des plantes cultivées en France FCU

Issu du projet VESPA, construit à la main à l'aide de plusieurs sources : le Larousse agricole, Wikipédia français, le registre parcellaire, les statistiques de l'agreste. Point de départ les céréales.

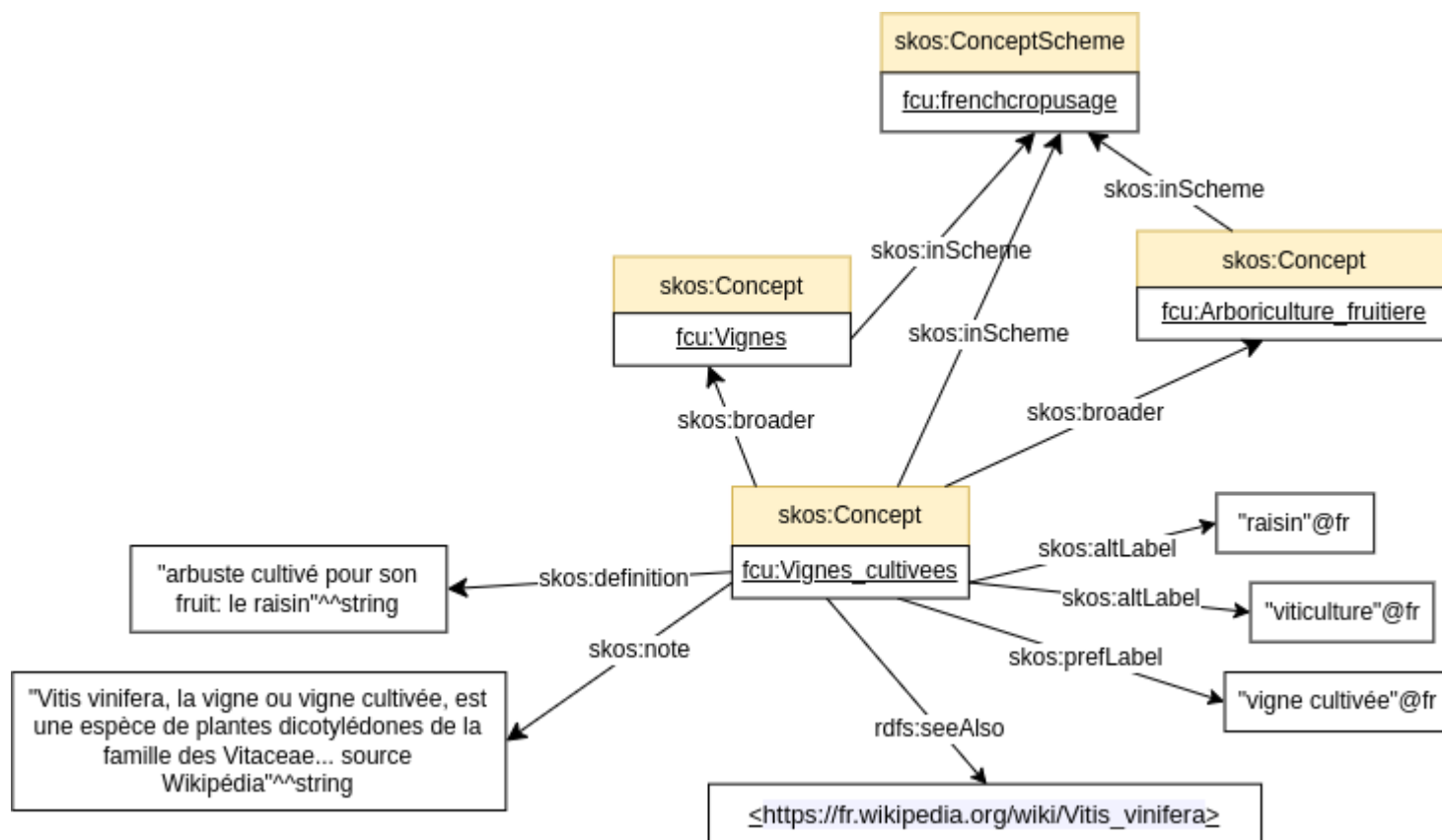
Licence ouverte, un DOI fourni par data.inrae, Disponible sur agroportal et sur le gitlab irstea <https://gitlab.irstea.fr/copain/frenchcropusage> et sur ontology.irstea.fr

Mise à jour pendant toute la durée du projet D2KAB : Version 2.1 et 2.2, la version 3 est disponible

- Maraîchage ; travail en commun avec Elzeard à l'aide de plusieurs nouvelles sources Bonduelle, FranceAgriMer, Encyclopédia Universalis, AM Nagelseim, la ferme du Bec Hellouin
- Fourrage : travail avec un expert à partir des fiches du GNIS
- Arboriculture / sylviculture : extractions de Anna Chepaikina, expertise acta
- Plante de service : travail en cours avec Elzeard, Acta et INRAE.
- Réutiliser dans la base GECCO, la publication sur le web de la base Ephy fait par D2KAB, la serre des savoirs d'Elzeard.



Thésaurus French Crop Usage FCU

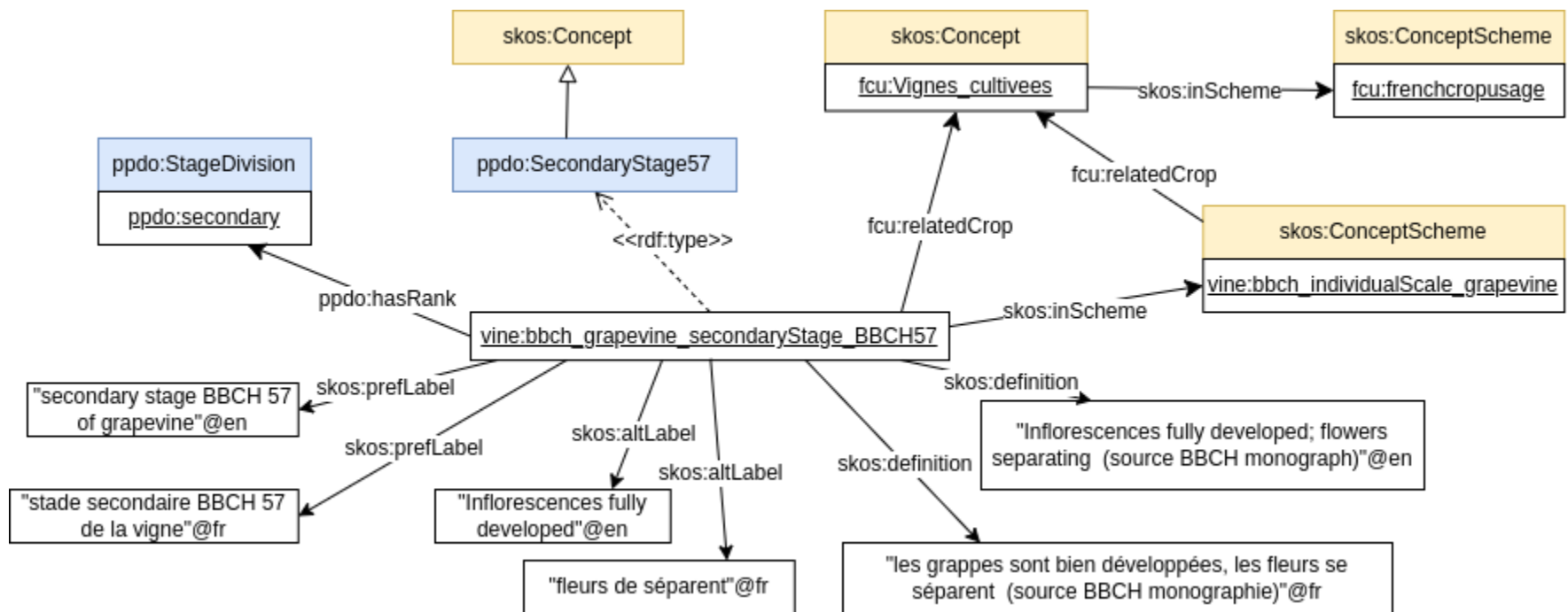


BBCH-based Plant Phenological Description Ontology

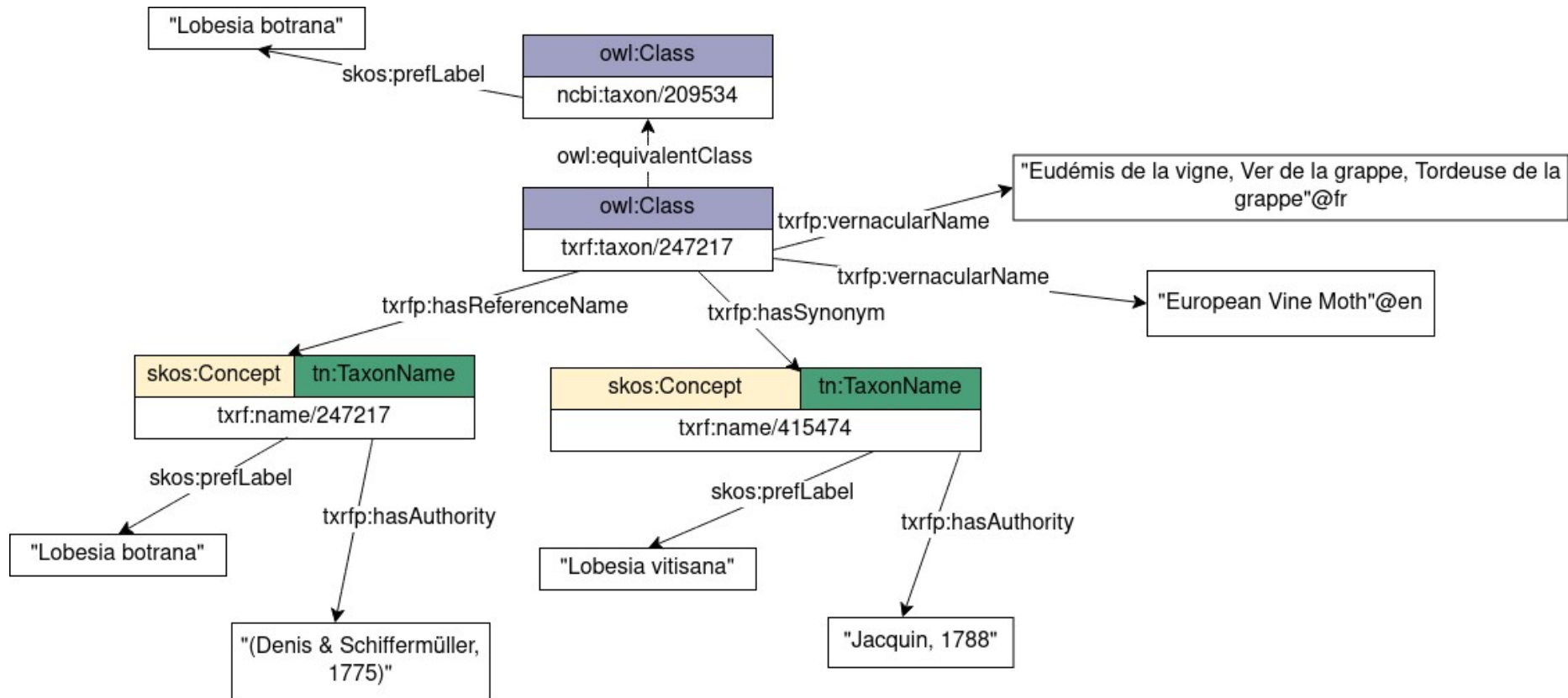
Basé sur la description des échelles phénologiques BBCH

Création d'un thesaurus (skos:ConceptScheme) par échelle phénologique.

Vignes 5 échelles phénologiques ont été alignées : EL, BBCH grapevine, Baggiolini, IFV Label et IFV Epicure

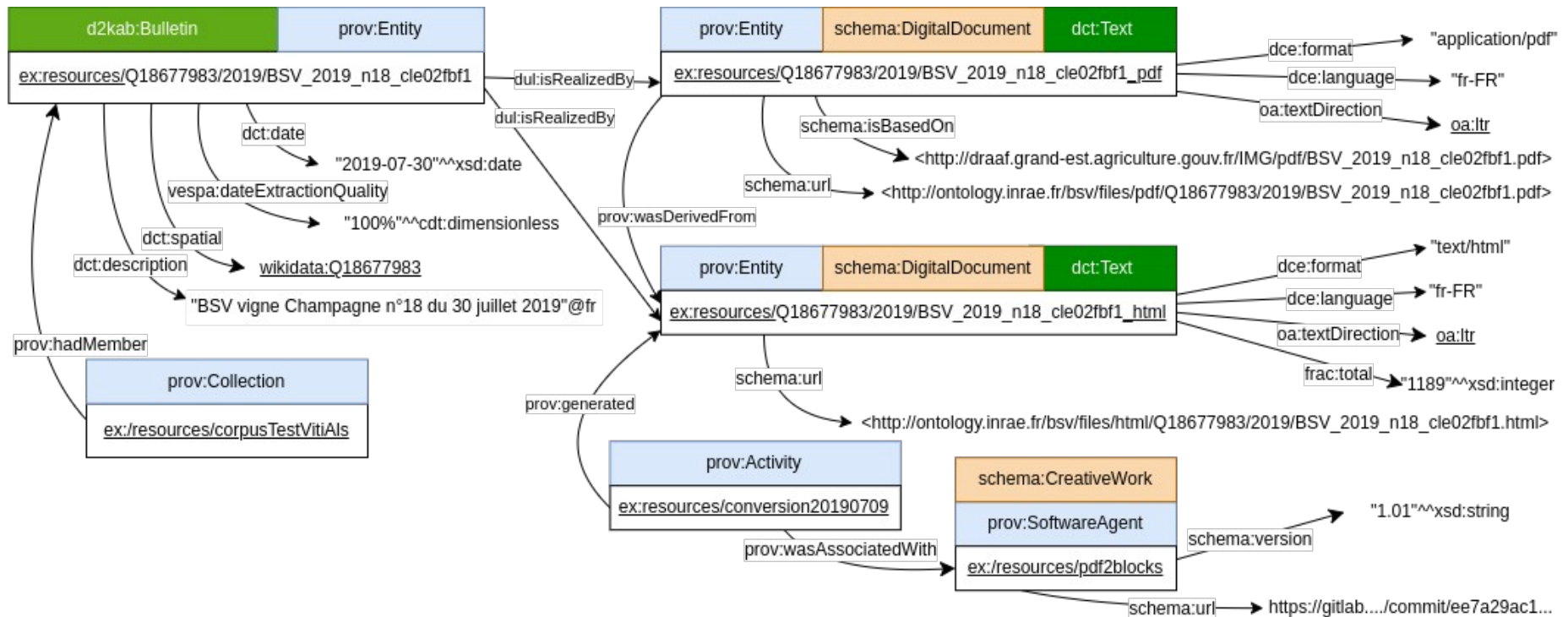


Taxonomie scientifique: TAXREF-LD et NCBI Taxonomie



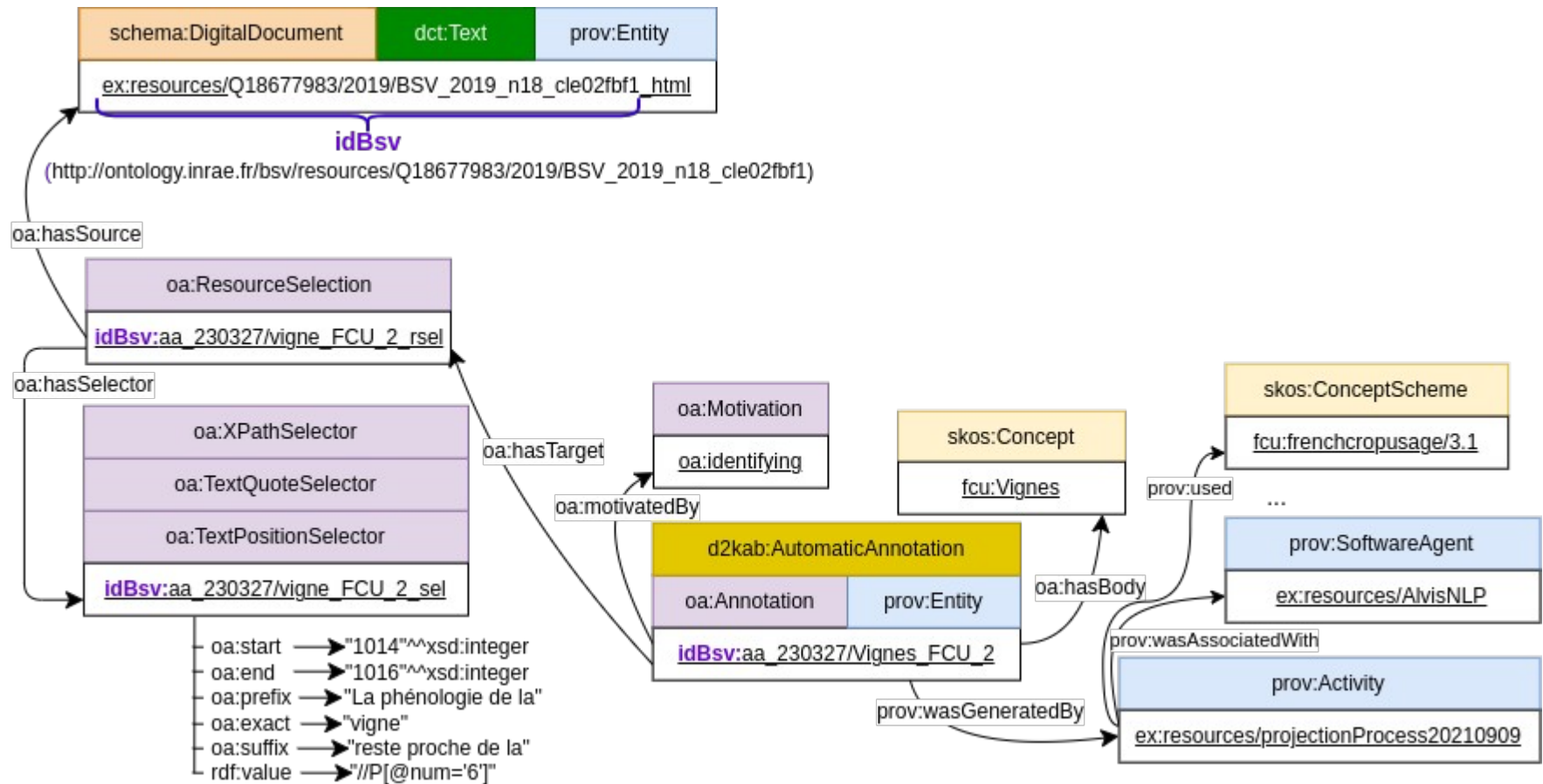
Structure des données

Représentation des BSV



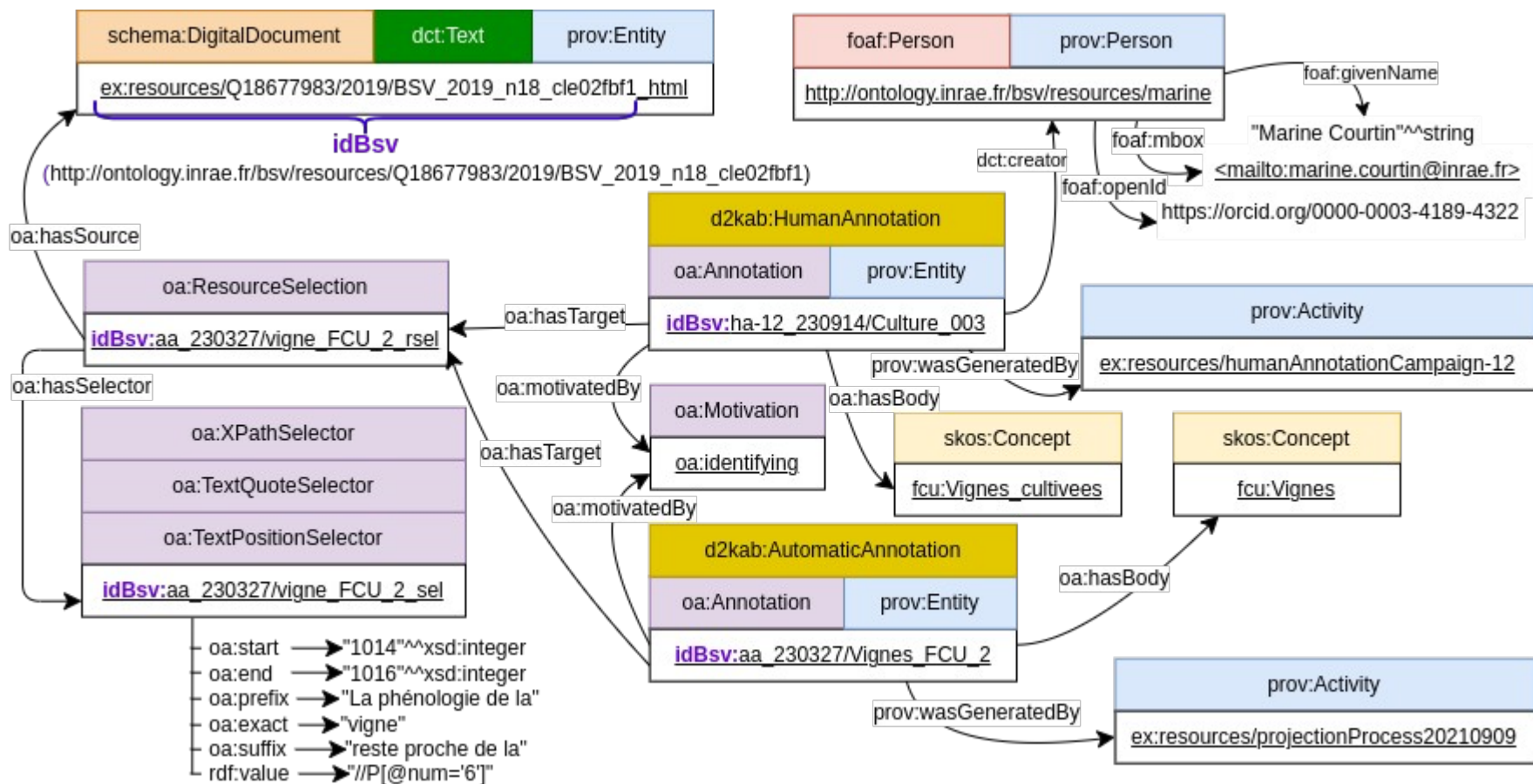
Structure des données

Annotations automatiques



Structure des données

Annotations automatiques et annotations manuelles



Liens vers les ressources

- NCBI taxonomy : <https://bioportal.bioontology.org/ontologies/NCBITAXON/>
- TAXREF-LD : <https://doi.org/10.5281/zenodo.6940891>
- Vitioeno : https://vitioeno.mistea.inrae.fr/resource/app/germplasm?rdf_type=vocabulary%3AVariety
- PPDO : <https://agroportal.lirmm.fr/ontologies/PPDO>
- FCU : <https://agroportal.lirmm.fr/ontologies/CROPUSAGE>
- Wikidata : <https://query.wikidata.org>



Conclusion

- Les ontologies du web sont des schémas de données FAIR qui contiennent des patrons de conception (des bonnes pratiques de modélisation)
- Les ontologies ont un cout: elles sont difficiles à comprendre et à réutiliser correctement
 - Quel informaticien écrit la documentation de son logiciel ou de sa base de données?
 - Qui lit les notices d'utilisation des appareils ou des médicaments?
- Les technologies du Web Sémantique changent les paradigmes de modelisation: base de graphes, Property First Class Citizen, Open World Assumption, ...



Travaux Pratiques

Calculer des indicateurs de changements climatiques

- Un graphe de connaissances des mesures météorologiques d'une station agricole de INRAE dupliqué sur deux serveurs
- Découverte du graphe et des différentes ontologies associées don't SOSA
- Interrogation du graphe pour calculer des indicateurs climatiques
- Le TP est visible sur <https://ontology.inrae.fr/pmwiki.php/Site/WeatherIndicators>

