



Séminaire Open Data en Agriculture

Technologies web de données pour la publication de données météo

Catherine ROUSSEY

équipe Copain

Laboratoire TSCF

Pour mieux
affirmer
ses missions,
le Cemagref
devient Irstea



www.irstea.fr

07 Janvier 2015 à Paris



Plan

Problématique: hétérogénéité des données capteurs

Quelques exemples

La station météorologique de la ferme de Montoldre

Publication de données météo de l'Irstea sur le web de données

L'ontologie Semantic Sensor Network du W3C

Quelques exemples

Notre réseau d'ontologies

Schéma d'une observation

Workflow de génération des données RDF à partir de CSV

Conclusion et Perspectives



Introduction

IRSTEA UN FOURNISSEUR DE DONNÉES

Irstea possède des réseaux de capteurs pour observer une variété de phénomènes naturels: météo, hydro, pollution etc...

Comment rendre accessible ces données ?
au personnel de l'Irstea mais pas que...

Le site web BDOH permet de rechercher les fichiers de données capteurs des observatoires de l'Irstea

Un premier essai de publication de données de capteurs avec les technologies web sémantique: données de l'Irstea sur le web de données



Problématique des données capteurs

HÉTÉROGÉNÉITÉ DES DONNÉES MÉTÉO

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

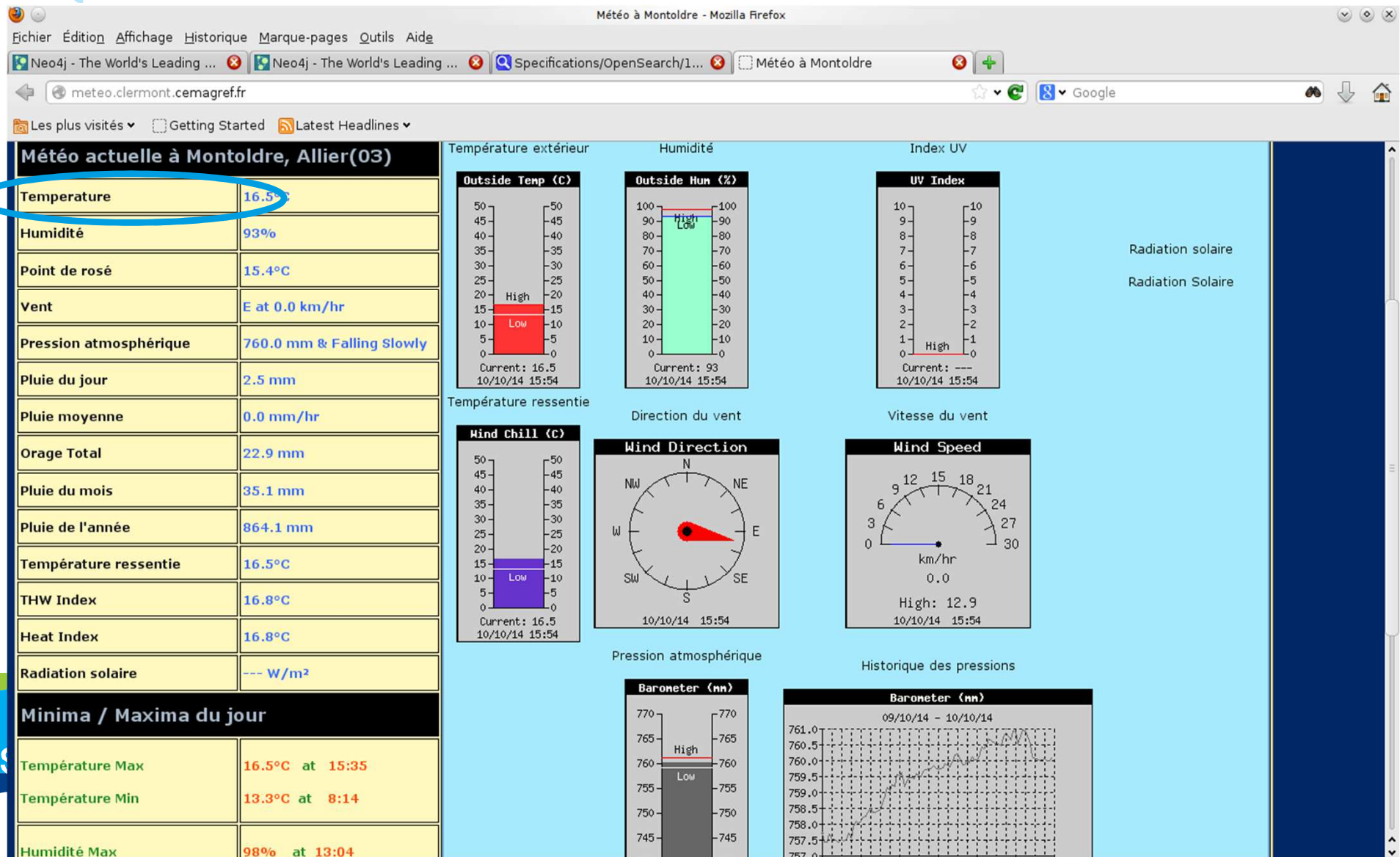
Problématique des données capteurs

DIFFÉRENTS TYPES DE CAPTEURS



Problématique des données capteurs

STATION MÉTÉOROLOGIQUE DE MONTOLDRE



Measurement Location: wan4, Wannengrat

7

Deployment/Fieldsite Information [hide]

Deployment Wannengrat (link: [Home](#))

Measurement Location wan4

Coordinates 46.803716666° N, 9.778030555° E

Altitude 2417 m

Swiss Northings

Swiss Eastings

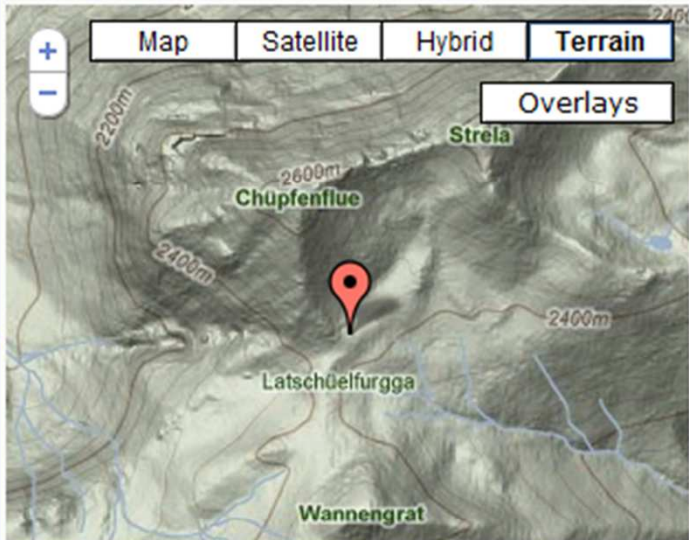
Experiments: Measuring and Modelling the Snow Distribution on Wannengrat ([about](#))

Edit information [Click here](#)

Tasks

- ▶ [Register an observation](#)
- ▶ [Register a sensor deployment](#)
- ▶ [Bulk Load information](#)

Location



Image



The v
locat

 Model	 S/N	 DB name	 Link
200-27005 UVW	2100	wind_speed_northerly_component wind_speed_westerly_component wind_speed_vertical_component	4da160f
SR50	C5635	snow_height	A948534
Hygroclip S3	45614	<u>air_temperature</u> relative_humidity	E57d1e4

Other Information [hide]

Source JP Calbimonte



GSN Server :: GSN

HOME DATA MAP FULLMAP

Auto-refresh every :

wan4 03/08/2009 11:10:00 +0200

Real-Time	Addressing	Structure	Description	Download
record 7920.0 relative_humidity 74.44 air_temperature 7.326 wind_speed_northerly_component 8.13 wind_speed_westerly_component -10.28 wind_speed_vertical_component -10.04 snow_height -198.3				

wannengrat_unterhalb_felsen 03/06/2010 02:06:00 +0200

Real-Time	Addressing	Structure	Description	Download
wind_direction null wind_speed null solar_rad null snow_water_content_1 null snow_water_content_2 null snow_temp_1 null snow_temp_2 null air_humid null air_temp null air_temp_ir null snow_surface_temp_ir null				



Problématique des données capteurs

POURQUOI PUBLIER LES DONNÉES MÉTÉO SUR LE WEB DE DONNÉES ?

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)



Publication de données météo de l'Irstea sur le web de données

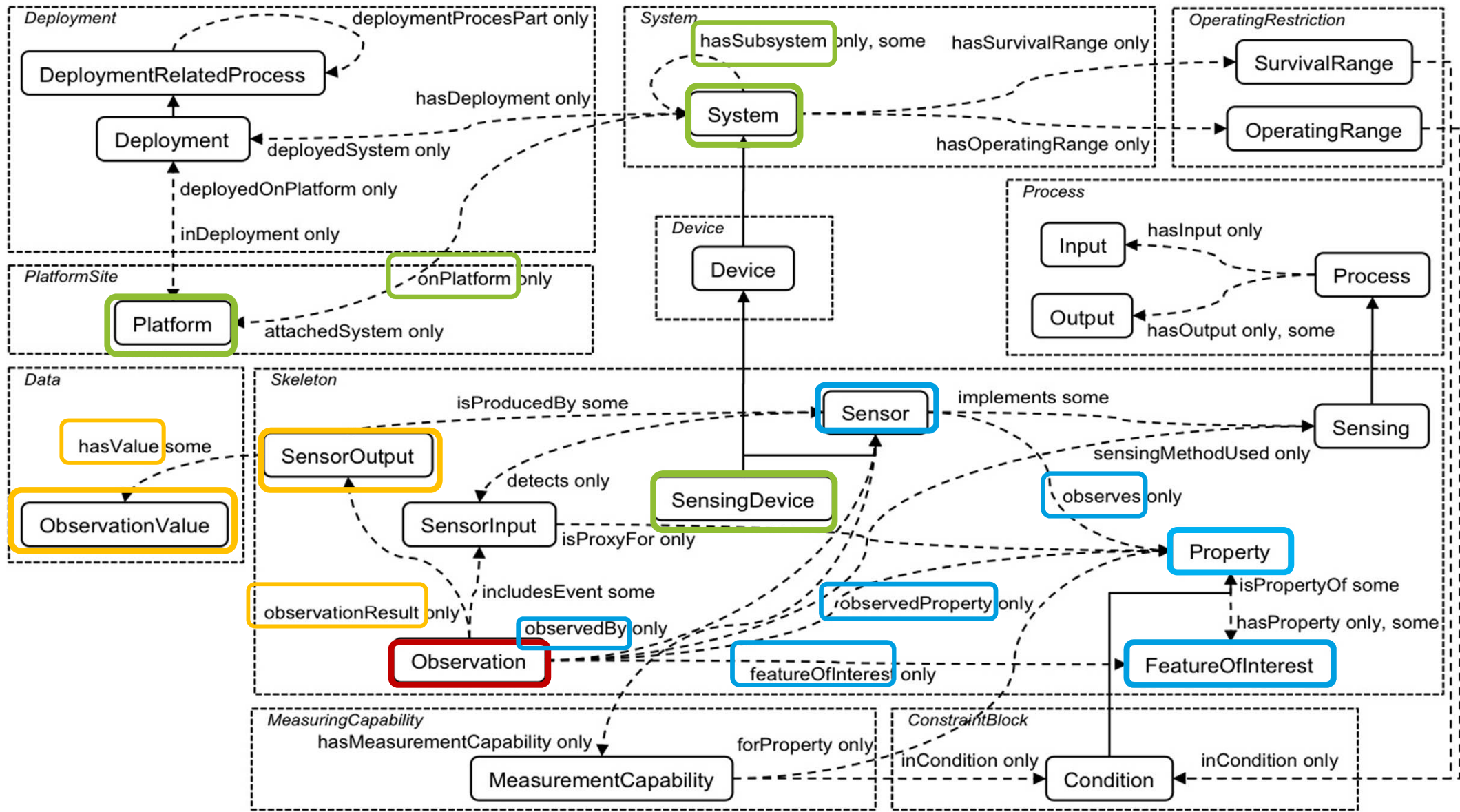
Pour mieux
affirmer
ses missions,
le Cemagref
devient Irstea



www.irstea.fr

Semantic Sensor Network Ontology

EN COURS DE NORMALISATION





Exemples d'ontologies associées à SSN

PUBLICATION DE DONNÉES MÉTÉO

Feature of interest, spatial, time

- AEMET (Agencia Estatal de Meteorologia)
AEMET, WGS84, Geobuddies, W3C Time
- Swiss Experiment project
SWEET, WGS84, QUDT
- ACORN-SAT (Australian Bureau of Meteorology)
WGS84, UK Intervals, DUL, Data Cube
- SMEAR (Finnish Station for Measuring Ecosystem Atmosphere Relations)
SWEET, Geoname, WGS84, DUL, Data Cube, Situation Theory



Notre réseau d'ontologies

Semantic Sensor Network : le squelette

Sensing Device

Ontology for meteorological sensor (aws)

Feature of Interest

Climate and Forecast (cf-feature + cf-property)

Platform location

GeoSPARQL and Location Core Vocabulary (geosparql + locn)

Observation

W3C Time Ontology (time)

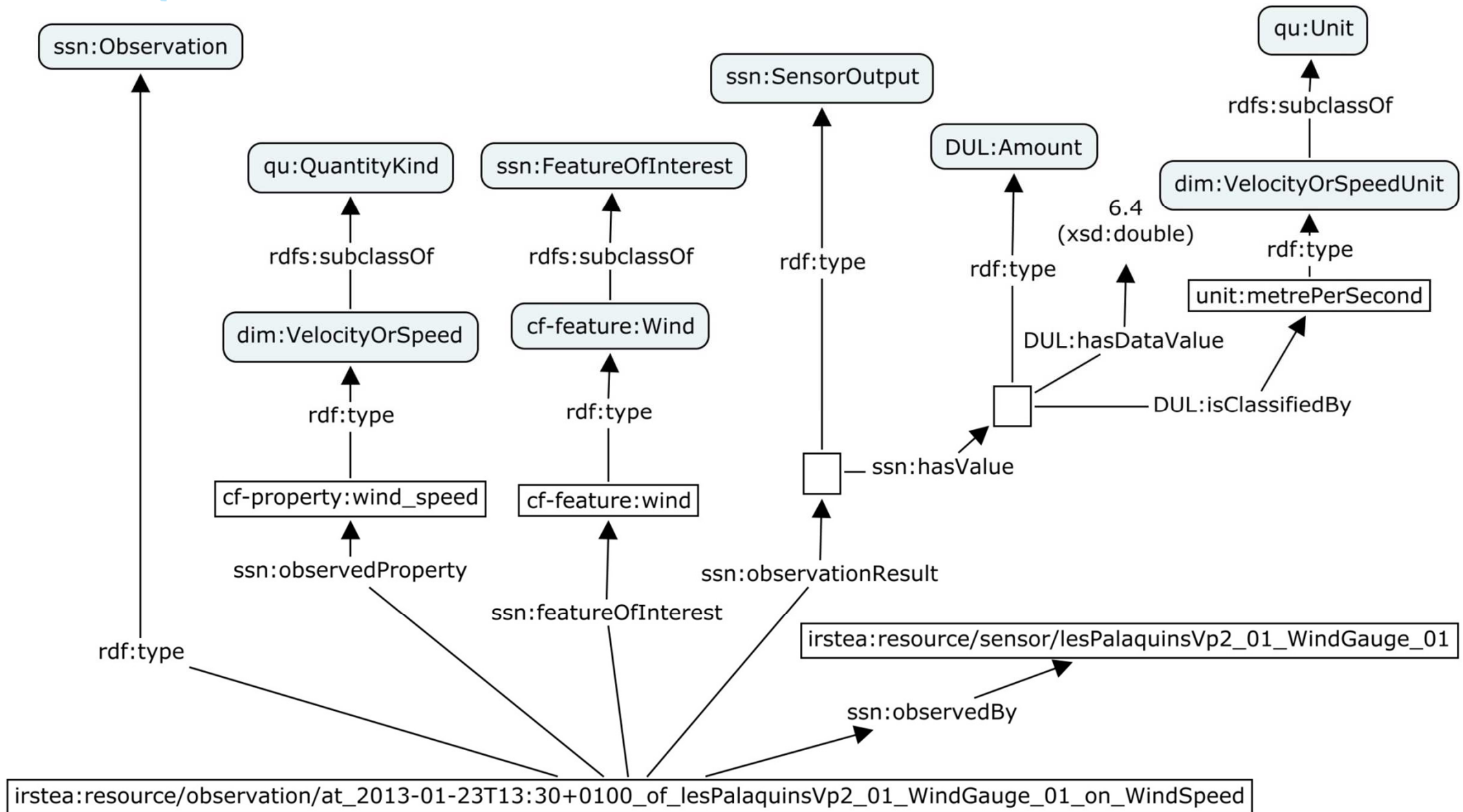
Observation value

Library of Quantity Kind and Units (qu + dim)

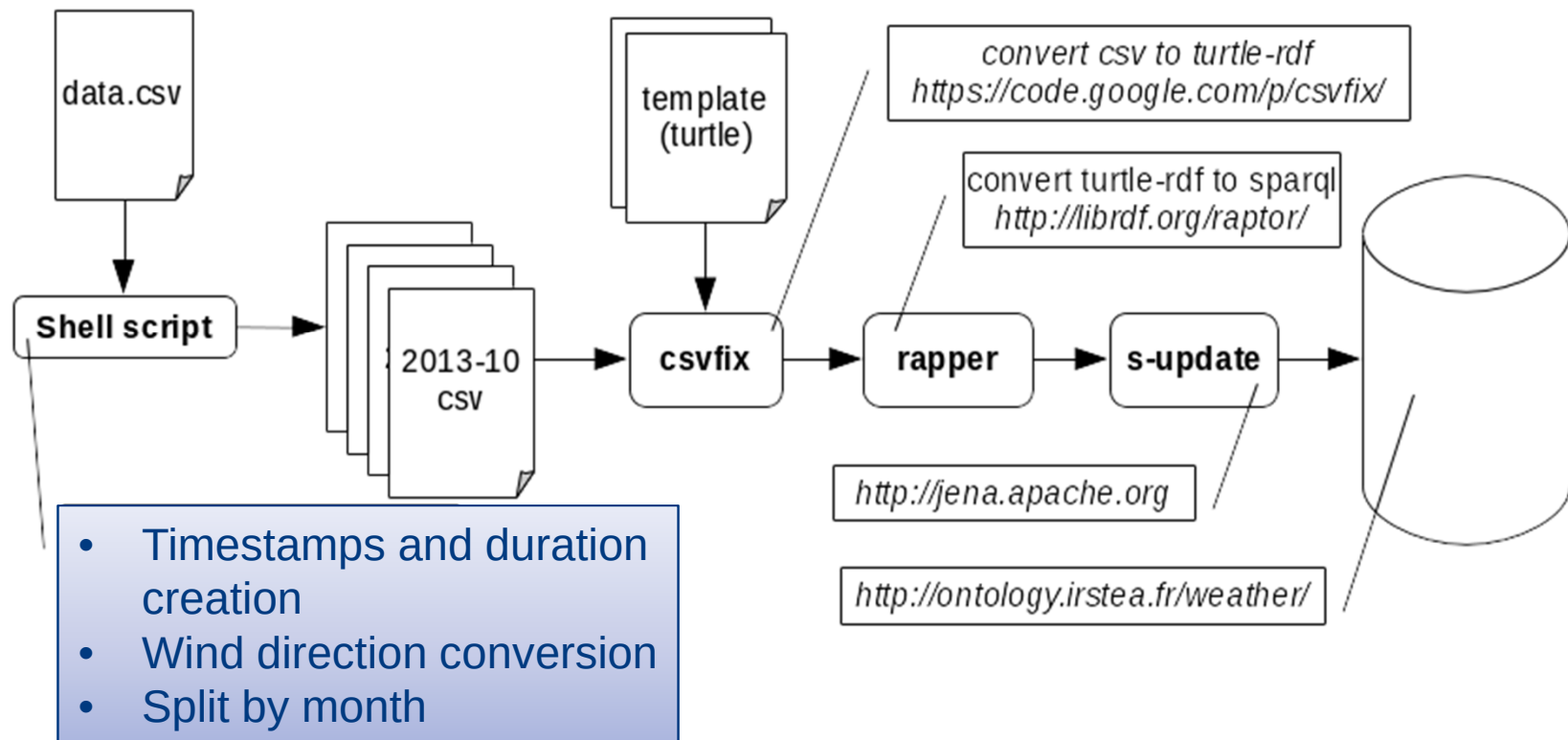
Dolce Ultra Light (dul)

Description d'une "Observation"

EXEMPLE D'UN ANÉMOMÈTRE



Transformation de CSV pour la publication



Publication des données de la station météo

Données publiées sous la licence open data etalab

SPARQL endpoint visible sur datahub

<http://ontology.irstea.fr/weather/snorql/>

Jeux de données

8 types de mesure * 48 mesures/jour * 365 jours * 4 ans =
560 640 observations

Requete SPARQL

```
PREFIX ssn:<http://purl.oclc.org/NET/ssnx/ssn#>
PREFIX cf-property:<http://purl.oclc.org/NET/ssnx/cf/cf-property#>
PREFIX dul:<http://www.loa-cnr.it/ontologies/DUL.owl#>
SELECT DISTINCT ?o ?data ?unit
WHERE {
  ?o a ssn:Observation.
    ?o ssn:observedProperty cf-property:air_temperature.
    ?o ssn:observationResult ?output.
    ?output ssn:hasValue ?value.
    ?value dul:hasDataValue ?data.
    ?value dul:isClassifiedBy ?unit.
} LIMIT 20
```



Conclusion et Perspectives

Publication des données de la station météo (2014 à jour)
pas si simple au début...

Publication d'autres données capteurs

Proposer une interface d'interrogation en français plus intuitive qu'une requête SPARQL

Lier nos données aux données de l'IGN pour une interrogation spatialisée

Tester différents serveurs RDF (requête avec agrégats)

Découverte et utilisation de ces données ?

Utiliser ces données dans des modèles météorologiques

Conversion au format netcdf?



Merci de votre attention

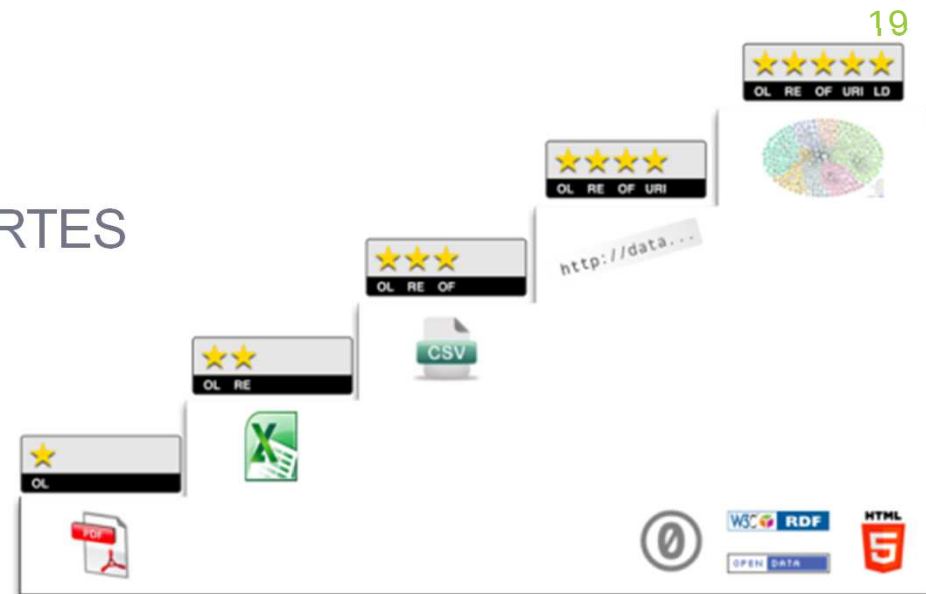
Pour mieux
affirmer
ses missions,
le Cemagref
devient Irstea



www.irstea.fr

Modèle 5 étoiles

WEB DE DONNÉES LIÉES OUVERTES



★ 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

Linked Open Data (LOD)

WEB DE DONNÉES LIÉES OUVERTES

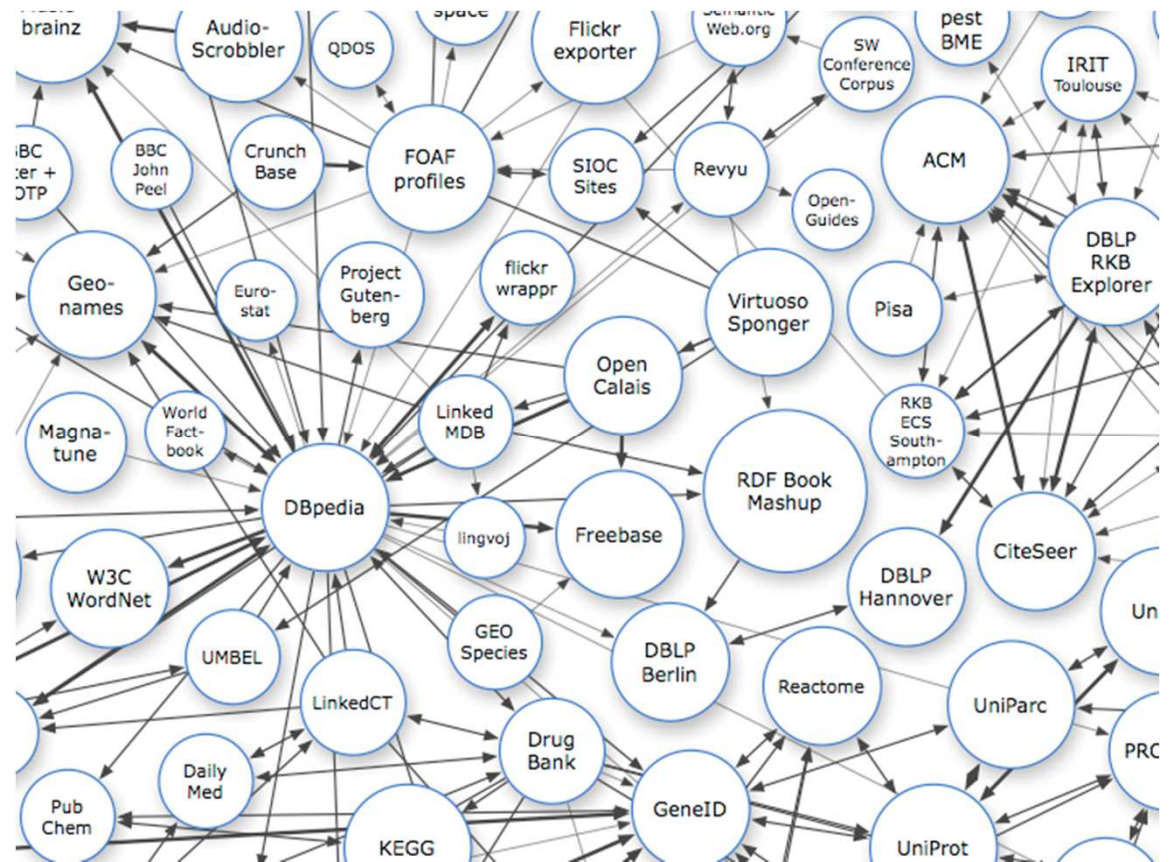
An extension of the current Web...

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

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

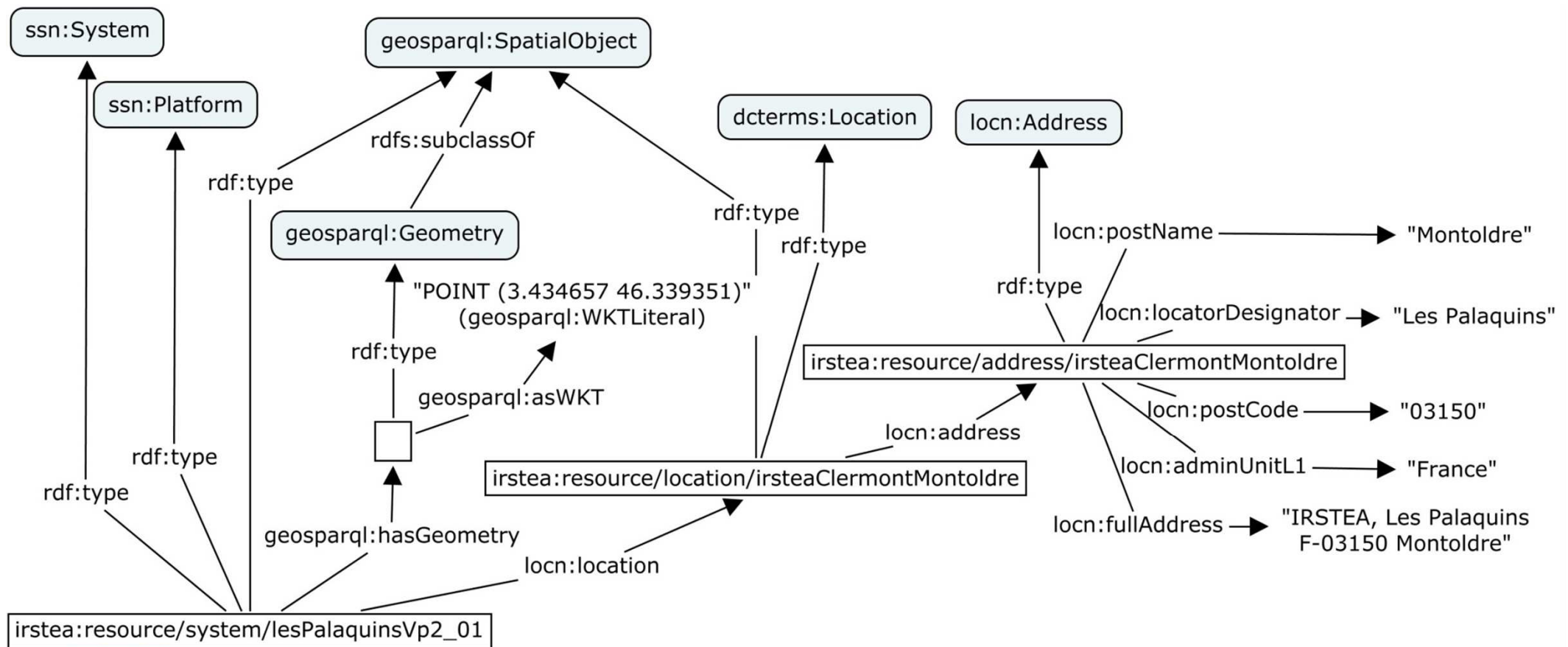
... better enabling them to work in cooperation

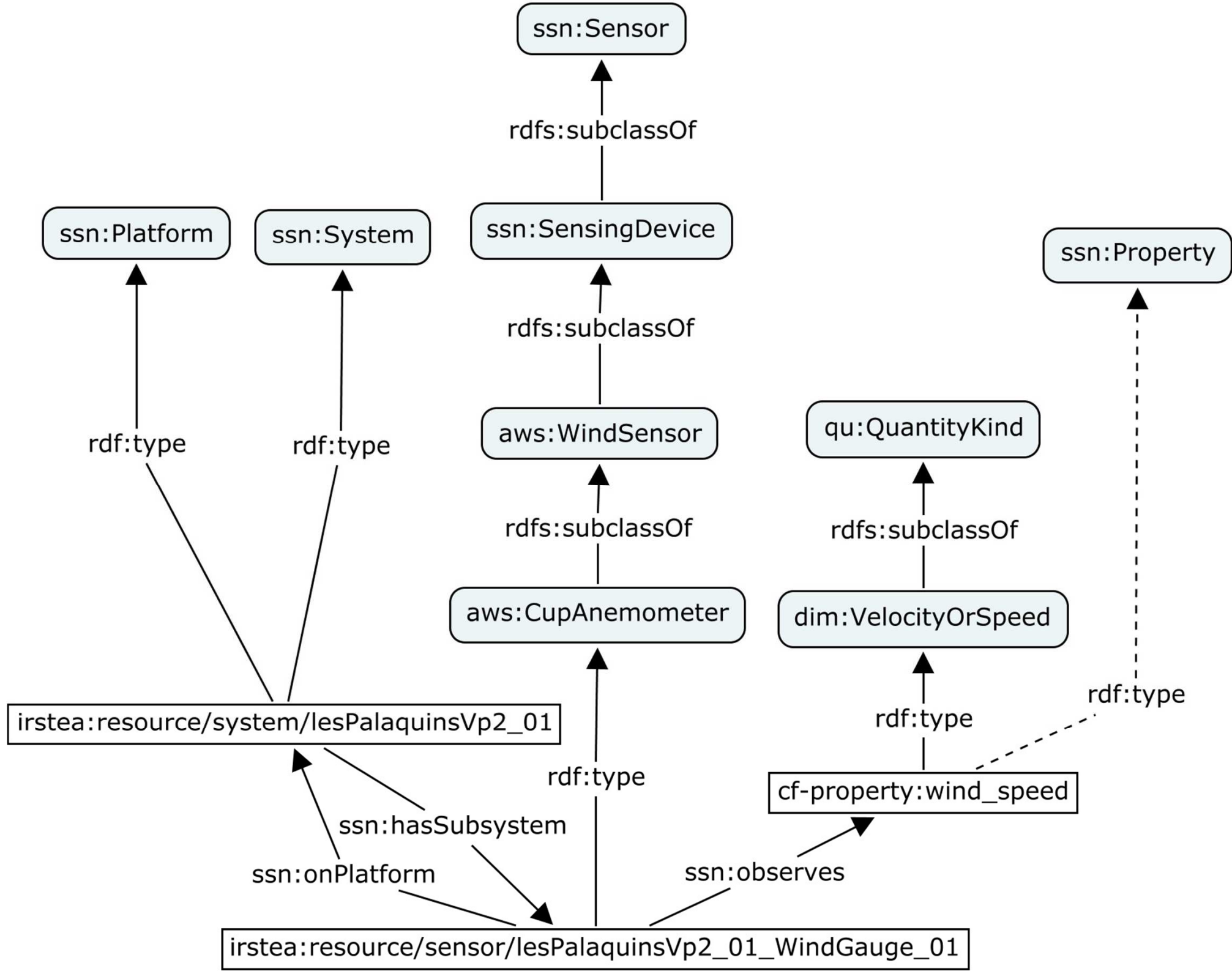
And clear principles on how to publish data



Description of Weather Station

SSN + LOCATION + GEOMETRY





Time Instant Observation

