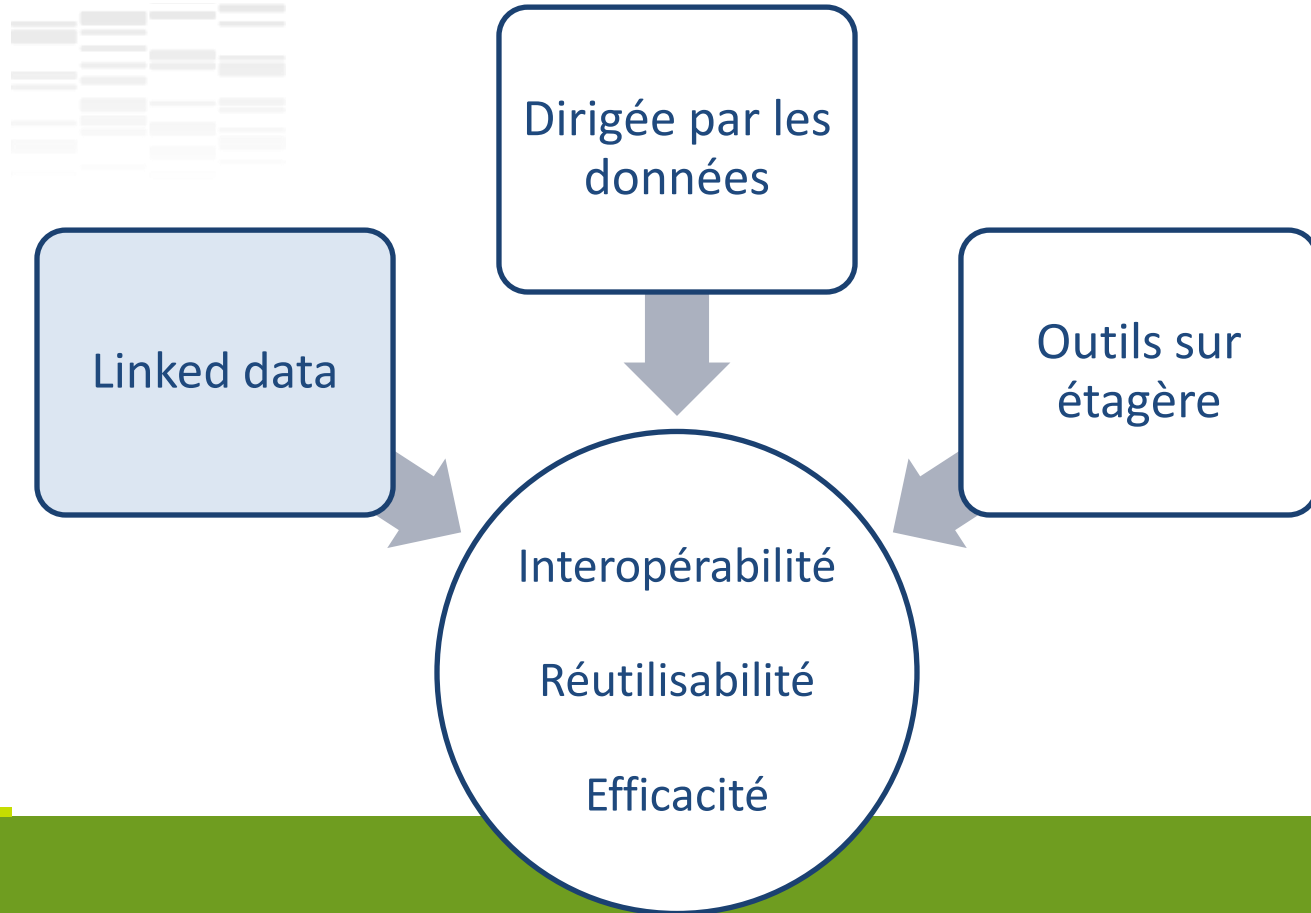




Publication de données d'observation dans le Web de données



Approche



Pourquoi publier dans le Web de données

- ❖ **Une voie vers l'intégration des données en interne**
- ❖ **L'opportunité de s'insérer dans un écosystème**
- ❖ Un moyen supplémentaire de valoriser l'expertise des chercheurs
- ❖ Une manière intelligente de répondre à un impératif : l'ouverture des données de la recherche
- ❖ Une occasion de revisiter la qualité des données produites : clarification, harmonisation, enrichissement, explicitation et documentation des données de manière standardisée

Publier dans le Web de données

Principes des Linked Data : W3C **Linked Data Glossary**

1. Utiliser des URIs pour nommer les choses
2. Utiliser des URIs accessibles via HTTP

<http://opendata-preprod.inra.fr:8080/public/Chenilles/CampagneObservation>

3. Donner l'accès aux données utiles en utilisant des standards ouverts comme
RDF, SPARQL
4. Lier ses données à d'autres données en utilisant leurs URIs

Publier dans le Web de données

Princip

❖ Utilis

❖ Utilis

<http://opendata-preprod.inra.fr:8080/public/page/Chenilles/CampagneObservation>

❖ Don

RDF

❖ Lier

Campagne observation | INP x

opendata-preprod.inra.fr:8080/public/page/Chenilles/CampagneObservation

Courrier :: Bienvenue ... ShareThis Bookmarklet

Campagne observation at INRA

<http://opendata-preprod.inra.fr:8080/public/page/Chenilles/CampagneObservation>

Property	Value
rdfs:comment	(xsd:string)
is rdfs:domain of	<ul style="list-style-type: none">inra:anneeCampagneinra:especeObserveeinra:etudeRetenue<http://purl.org/dc/elements/1.1/date>
rdfs:label	Campagne observation (xsd:string)
is rdfs:range of	<ul style="list-style-type: none">inra:dansCampagneObsinra:participeAcampagne
rdfs:subClassOf	<http://rdf.geospecies.org/ont/geospecies#Project>
is rdf:type of	<ul style="list-style-type: none">inra:CampagneObs2011inra:CampagneObs2012
rdf:type	<http://www.w3.org/2002/07/owl#Class>

This page shows information obtained from the SPARQL endpoint at <http://localhost:8080/openrdf-sesame/repositories/test2>.
[As Turtle](#) | [As RDF/XML](#) | [Browse in Disco](#) | [Browse in Tabulator](#) | [Browse in OpenLink Browser](#)

Publier dans le Web de données

Princip

❖ Utilis

❖ Utilis

<http://o>

❖ Don

RDF

❖ Lier

Campagne observation | INRA x

opendata-preprod.inra.fr:8080/public/page/Chenilles/CampagneObservation

Courrier :: Bienvenue ... ShareThis Bookmarklet

Campagne observation at INRA

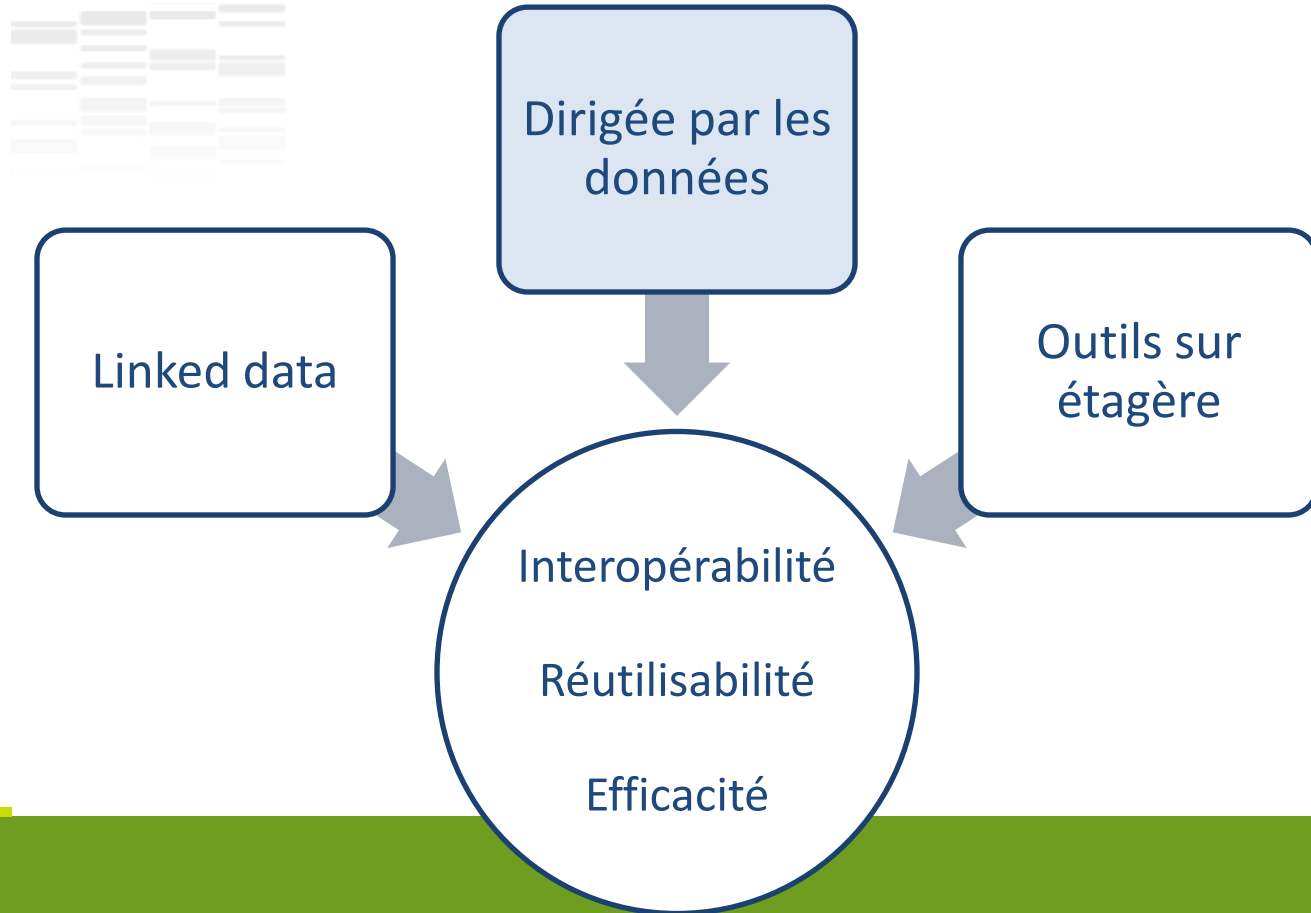
<http://opendata-preprod.inra.fr:8080/public/Chenilles/CampagneObservation>

Property	Value
rdfs:comment	(xsd:string)
is rdfs:domain of	<ul style="list-style-type: none">inra:anneeCampagneinra:especeObserveeinra:etudeRetenue<http://purl.org/dc/elements/1.1/
rdfs:label	Campagne observation (xsd:string)
is rdfs:range of	<ul style="list-style-type: none">inra:dansCampagneObsinra:participeAcampagne
rdfs:subClassOf	<http://rdf.geospecies.org/ontology
is rdf:type of	<ul style="list-style-type: none">inra:CampagneObs2011inra:CampagneObs2012
rdfs:type	<http://www.w3.org/2002/07/owl#Class>

This page shows information obtained from the SPARQL endpoint at <http://localhost:8080/openrdf-sesame/repositories/test2>.
[As Turtle](#) | [As RDF/XML](#) | [Browse in Disco](#) | [Browse in Tabulator](#) | [Browse in OpenLink Browser](#)

Annotations on the screenshot:
- A blue box labeled 'concept' points to the title 'Campagne observation at INRA'.
- A blue box labeled 'relation' points to the 'is rdfs:domain of' property and its values.
- A blue box labeled 'instance' points to the 'is rdf:type of' property and its values.

Approche



Comprendre et qualifier les données

- ❖ Récupérer toute documentation utile sur les données
 - ❖ Schéma (XML, base de données relationnelle), rapports, etc.
- ❖ Se procurer un échantillon des données
- ❖ Échanger avec les producteurs de données
- ❖ Désambiguïser, harmoniser, préciser

Le projet scientifique – ALTERPRO

INRA*, Plante&Cit , ONEMA, Plan Ecophyto 2018

- ❖ Contrôler les populations de Processionnaire du Pin
 - ❖ à un niveau tol rable permettant de prot ger hommes, plantes et animaux
 - ❖ avec des moyens  cologiques : pi ger les papillons m les avec des pi ges   ph romones pour limiter la reproduction
- ❖ Comparer l'efficacit  des pi ges et des ph romones disponibles sur le march 

* UE Entomologie et For t M diterran enne

Protocole de l'expérimentation

- 1 Pose des pièges à phéromones sur les sites test
- 2 Comptage des papillons piégés sur les sites test
- 3 Comptage des nids de chenilles sur les sites test et témoin

→ Période de piégeage des papillons

→ Période de dénombrement des nids de chenilles



Description des sites

Données géographiques

	A	B	C	D	E	F	G	H	I	J	K	L	M
1		Communes	Saint-Fargeau-Ponthierry	Obernai	Angers							Saint-Martin-de-Ré	Marenne
2	Description Sites Témoins	Type de configuration	alignement	pas de site témoin déterminé	îlot boisé							îlot boisé	îlot boisé
3		Description des sites	parc de loisir		Echangeur de la Baumette							Parcelle forestière	à proximité la piscine Marenne
4		Géolocalisation	48°31'56.85"N, 2°33'47.42"E		47°27'53.61"N, 0°34'14.82"O							46°15'3.24"N, 1°31'0.58"O	45°49'28.2"O, 1° 8'19.54
5		Surface	56 m		25100 m²							0,13 ha	10 ha
6		Historique du site	non communiqué		non communiqué							Comptage des nids par l'ONF depuis 2003 (données transmises)	non communiqué
7		Nombre de pins et/ou de cèdres	20		62							116	non communiqué
8	Sites Tests	Nombre de sites tests	1	1	7							1	1
9		Type de configuration	alignement	îlot boisé	alignement	îlot boisé	alignement	alignement	alignement	îlot boisé	alignement	îlot boisé	îlot boisé
10		Description des sites	parc de loisir	jardin privé	Roseaie I bord de route	Roseaie II Parc	Roseaie III butte antibruit	Roseaie IV quartier résidentiel	Arboretum V Falun d'Orgemont	Arboretum VI cours d'école	Arboretum VII Plaine de jeux	parc	bosquet
11		Géolocalisation	48°31'59.30"N, 2°33'29.47"E	48°27'34.36"N, 7°30'18.04"E	47°26'55.62"N, 0°34'21.40"O	47°26'48.09"N, 0°33'51.29"O	47°26'33.64"N, 0°33'28.63"O	47°26'38.59"N, 0°33'14.78"O	47°26'43.17"N, 0°32'29.08"O	47°27'10.16"N, 0°32'47.32"O	47°27'45.99"N, 0°31'39.55"O	46°10'54.61"N, 1°22'26.48"O	45°49'14.1"O, 1° 8'10.14
12		Surface	500 m	< 15 ares	3000 m²	17939 m²	8650 m²	8650 m²	3250 m²	5270 m²	1077 m²	4,6 ha	1 ha

Protocole

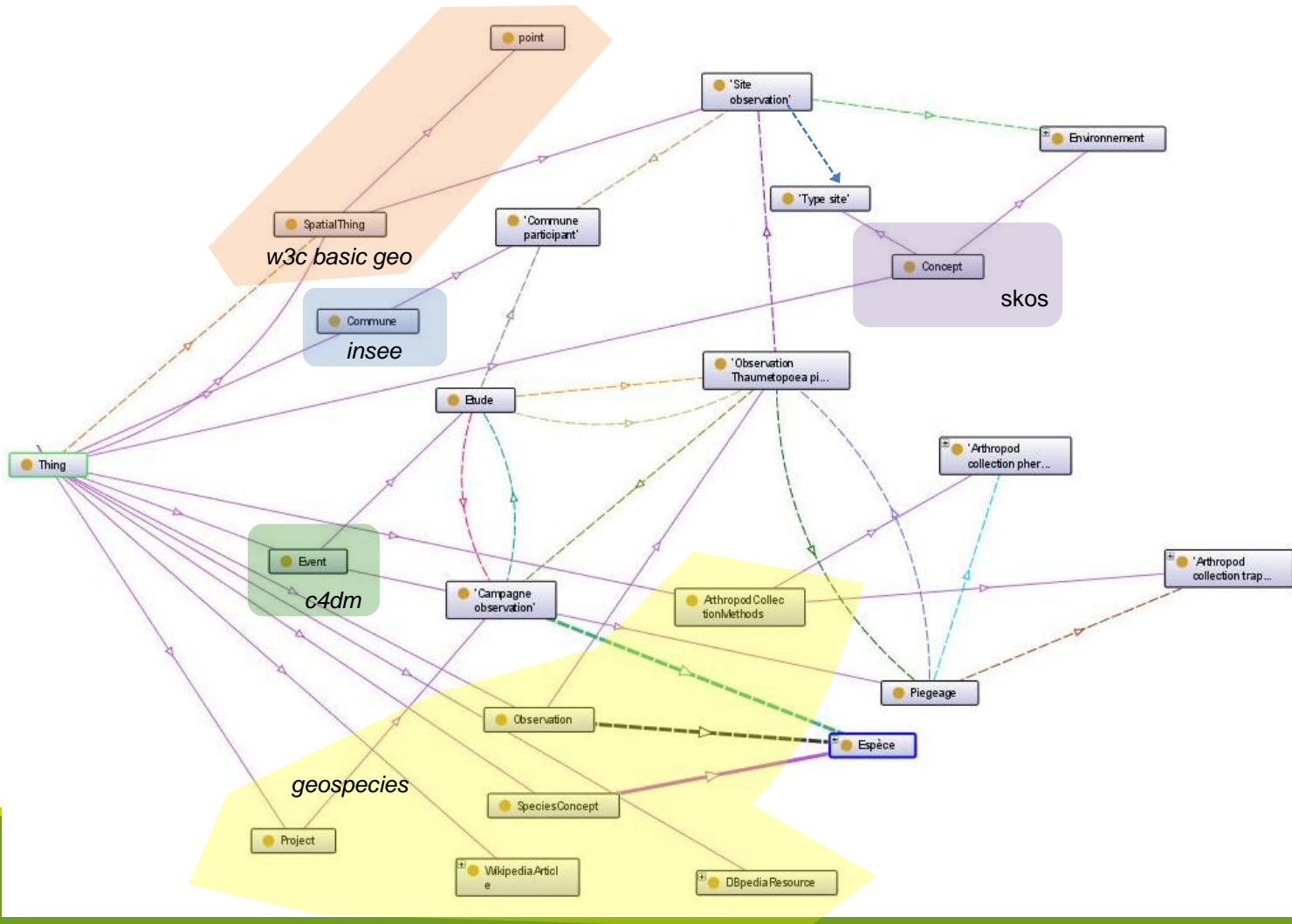
Relevés

Dispositifs pièges à phéromones été 2011

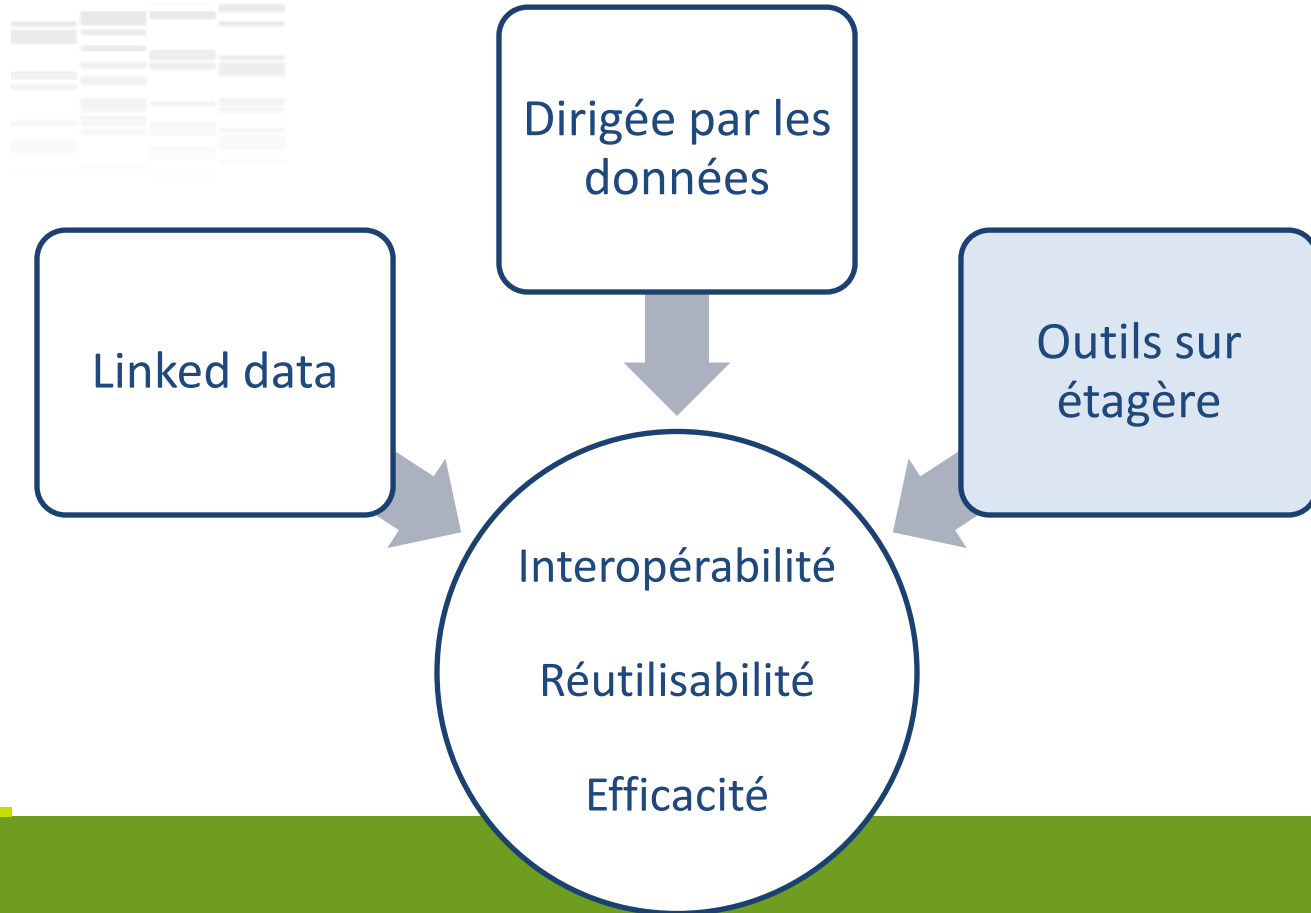
Relevés des nids réalisés à l'hiver 2011	non réalisés	non réalisés	non réalisés								non réalisés tém
	10	15	14 a) 1 b)	9	12 a) 1 b)	12 a) 2 b)	8 a) 1 b)	6 a) 1 b)	8	18	
Type de pièges	Procerex	Mastrap	a) Mastrap b) Nufarm	Mastrap	a) Mastrap b) Nufarm	a) Mastrap b) Nufarm	a) Mastrap b) Nufarm	a) Mastrap b) Nufarm	a) Mastrap b) Nufarm	Mastrap	Mast
Type de phéromones	Procerex	Procerex	Procerex	Procerex	Procerex	Procerex	Procerex	Procerex	Procerex	Procerex	Proce
Date de la pose des pièges	25 juillet (sem 30)	01 juin (sem 22)	01 juillet (sem 26)				01 juillet (sem 26) et 06 juillet (sem 27)	31 mai (sem 22) et 19 juillet (sem 29)	19 juillet (sem 29)	19 juillet (sem 29)	12 ju (sem
Premiers vols de papillons constatés	non communiqué	Sem 26 et 27	sem 25								non re
Relevés intermédiaires réalisés	non réalisés	19 juin 4 juillet	non réalisés								30 ju (sem
Nombre de papillons piégés au total	non communiqué	8	340	305	430	201	390	285	155	58	

Modéliser

- ❖ Une ontologie légère (OWL-Lite)
- ❖ Recours à des vocabulaires standards: RDF, RDFS, DC, Skos,
 - ❖ L'utilisation de **vocabulaires standards** accroît le potentiel **d'interopérabilité** et de **réutilisation** des données => LOV (Linked Open Vocabularies)
- ❖ Utilisation ou liens vers des concepts et des instances externes
 - ❖ Critères : qualité, fiabilité, pérennité
 - ❖ TaxonConcept, Geospecies, Event, FOAF, qudt,
 - ❖ Insee pour les communes



Approche

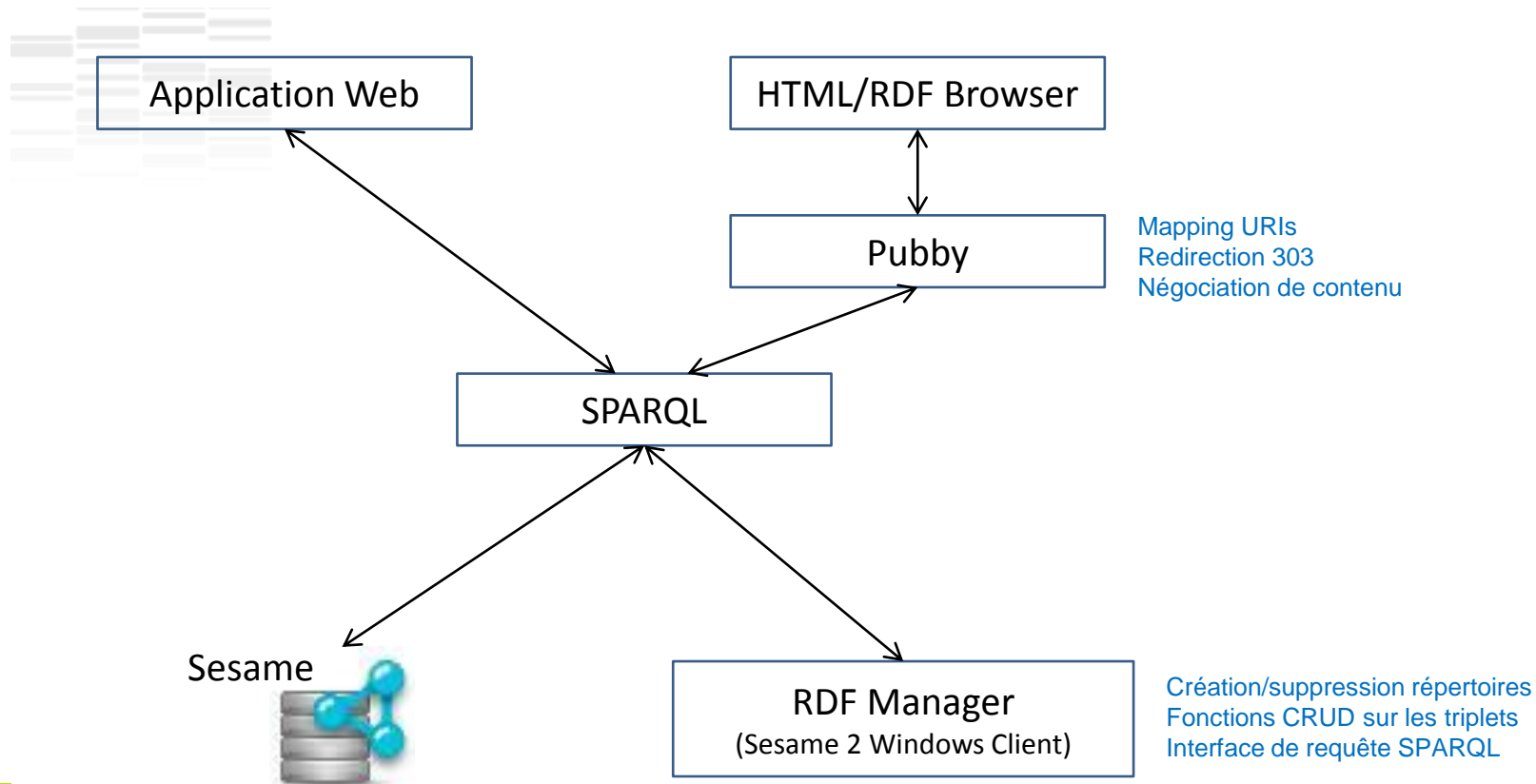


Créer la base de connaissances



- ❖ Utilisation d'URI (Uniform Resource Identifier) de type HTTP
- ❖ Transformation des données au format RDF
 - ❖ Par programme : XLWrap
 - ❖ Avec des outils : TopBraid, GoogleRefine
- ❖ TripleStore: Sesame
- ❖ Visualiseur RDF : Pubby

Infrastructure

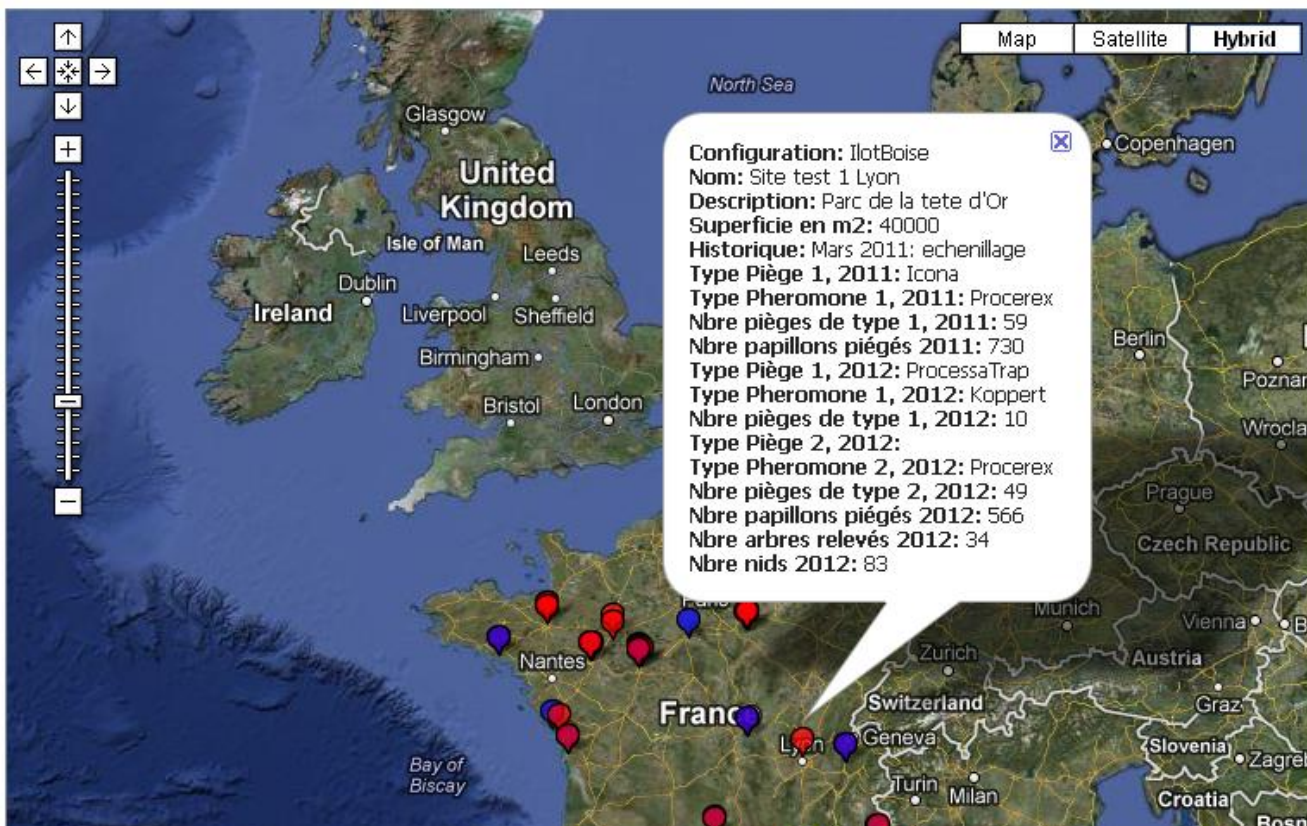


Sites d'observation

CARTE • DETAILS

69 SiteObservation

6 résultats sur 69 ne peuvent pas être tracés.



Search

Localisation

- 2 Aix-les-Bains
- 8 Angers
- 2 Arzon
- 4 Avignon
- 3 Bayonne
- 2 Entraunes

Configuration

- 1 (missing this field)
- 3
- 14 Alignement
- 3 ArbreSeul
- 48 IlotBoise

Type Site

- 23 Temoin
- 46 Test

get data file

Résultats

- ❖ Une ontologie et des données **réutilisables** :
 - ❖ vocabulaires connus et/ou documentés,
 - ❖ formats et protocoles standards
- ❖ Des données connectées et **interopérables**
- ❖ Une méthodologie et une chaîne de traitements pour être **efficaces** pour:
 - ❖ construire les ontologies
 - ❖ transformer et enrichir (lier) les données
 - ❖ publier les ontologies et les données

Compétences mobilisées

