

# Standardiser les données expérimentales pour faciliter l'innovation



Florence AMARDEILH, Baptiste DARNALA



Arnaud CHARLEROY, Catherine ROUSSEY,



Xavier DELPUECH,



Frédéric SALVI

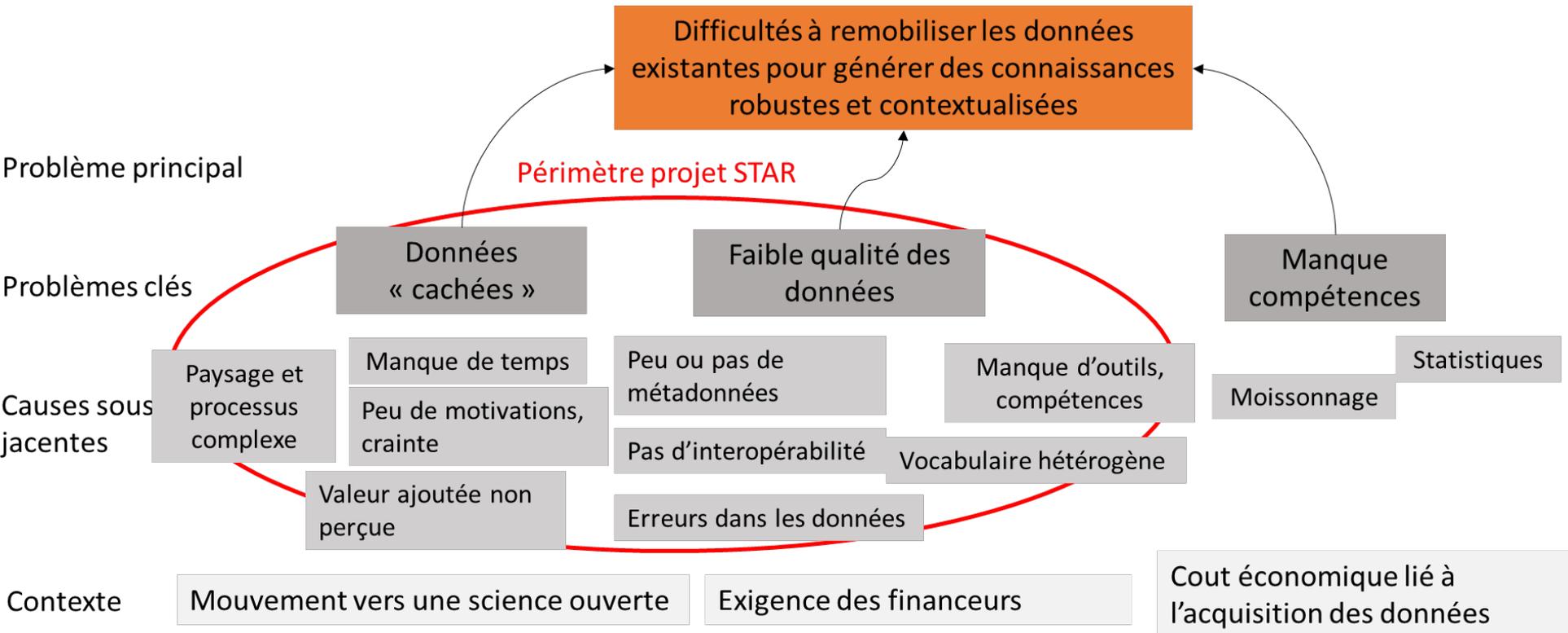


# Le projet STAR

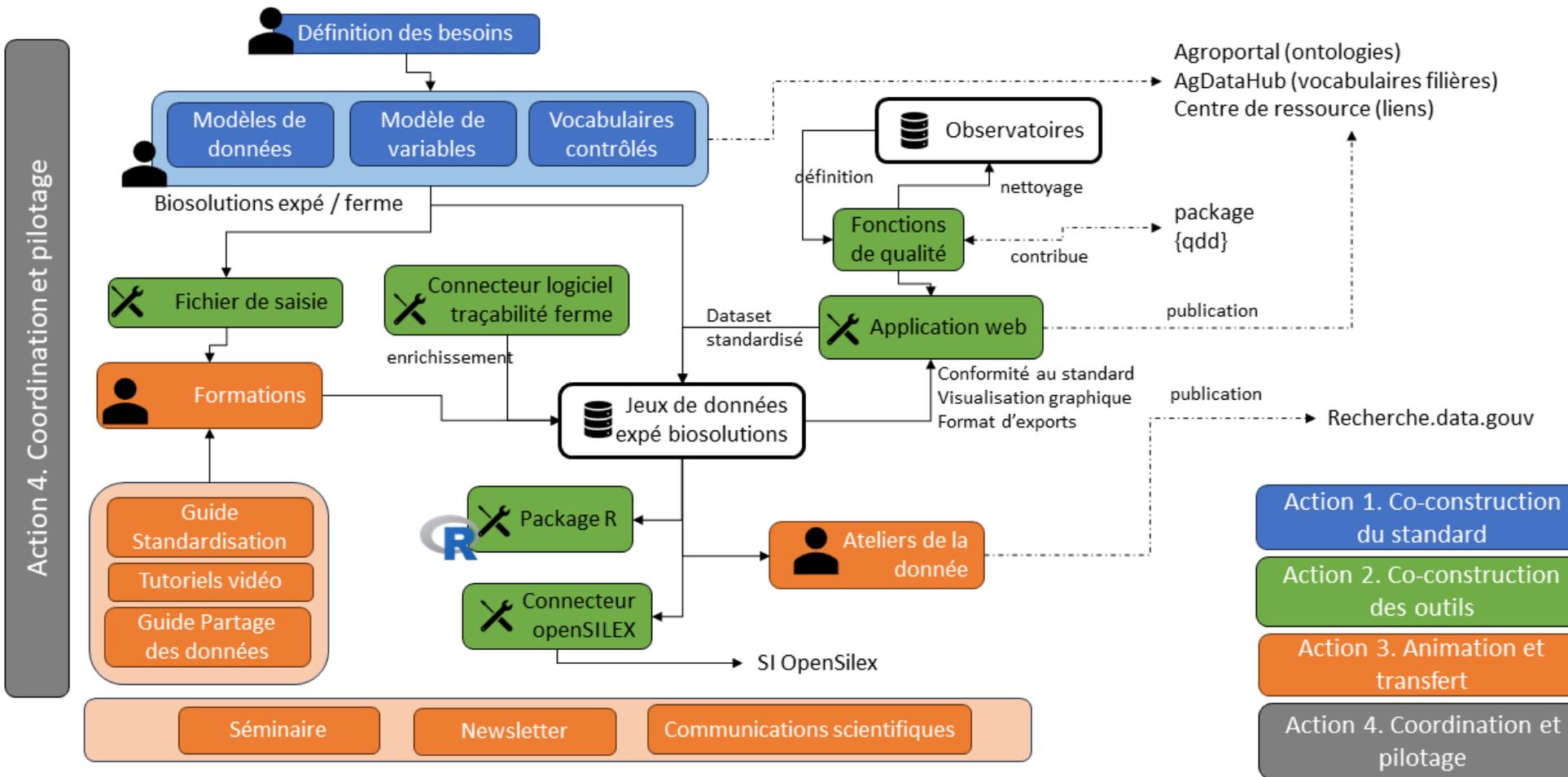
Standardiser les données expérimentales et techniques pour faciliter leur réutilisation et accélérer l'innovation et le développement agricole : application aux travaux sur les biosolutions

Accélérer l'innovation et l'adaptation des systèmes de productions agricoles est de plus en plus nécessaire dans un contexte de mutations accélérées par le changement climatique et la montée des enjeux environnementaux. C'est en particulier vrai pour le développement des **biosolutions** (biocontrôle et biostimulants), dont l'efficacité sur le terrain est parfois partielle et variable selon leurs conditions d'utilisation. Pour définir leur domaine d'utilisation, il est nécessaire à la fois de (i) **décrire finement les conditions de mises en œuvre** des produits (sol, climat, itinéraires techniques...) et (ii) **de mettre en commun des données** issues de situations de production variées. Le projet STAR a pour objectif de lever les freins à un partage et une mutualisation des données **expérimentales** (collectées par des expérimentateurs selon des protocoles définis sur des dispositifs contrôlés) et **techniques** (c'est-à-dire les données décrivant le ou les itinéraires techniques mis en œuvre sur l'exploitation). Pour cela, le projet vise à proposer un ou des **modèles de données** adaptables associés à des **vocabulaires contrôlés** pour obtenir des données interopérables. Le projet vise à coconstruire avec les expérimentateurs des **outils** simples et pratiques à mettre en œuvre **pour valider et produire des données standardisées**. Il inclut une action de **sensibilisation, formation et accompagnement** des acteurs.

# Le projet STAR : Arbre à problèmes



# Le projet STAR : actions



# Action 1 :

## Co-construction de ressources sémantiques pour l'interopérabilité des données

### Modèles de données

État de l'art et analyse des besoins sur les standards en expérimentation agricole  
Liste des questions de compétences  
Modèle des données d'expérimentation générique et adaptés  
Modèle viti (base startbox)  
Modèle de description des itinéraires techniques  
Modèle de variables : Thésaurus Terres Inovia, Thésaurus variables Vitis ontology enrichi

### Vocabulaires contrôlés

État de l'art sur les sources d'informations existantes (documents, glossaires, ...)  
Méthodes et outils pour la construction de vocabulaires terme : exemple les échelles phenologiques sont publiées à l'aide de PPDO  
Vocabulaires : cas d'usage Biosolutions, variables (dont méthodes...), autres

### Validation

Extrait des jeux de données formalisés et validés dans les différents modèles proposés avec les vocabulaires proposés

# Etat de l'art : modèle de données

## Expérimentations agronomiques

- Format d'échange de données ICASA
- Standard ISA spécification
- Minimal Information About Plant Phenotyping Experiment (MIAPPE)
- → Plant Phenotype Experiment Ontology (PPEO)
- Experimental Scientific Objects Ontology (OESO)
- Les ressources du réseau Crop Ontology : AgrO, Plant Ontology, ....

## Données agricoles

- Gestion des Informations de l'Exploitation Agricole (GIEA)
- Data Plot Sheet d'AGRO EDI Europe
- Crop Planning and Process Production Ontology (C3PO)



# Sources d'informations réutilisables

Trouver des définitions et du vocabulaire dans des documents de référence:

- European and Mediterranean Plant Protection Organization / Organisation Européenne et Méditerranéenne pour la Protection des Plantes. "PP1/152(4)- **Design and analysis of efficacy evaluation trials** in Bulletin OEPP/EPPPO, 2012, 42 (3), pp 367-381, ISSN 0250-8052. DOI: 10.1111/epp.2610, <https://pp1.eppo.int/standards/PP1-152-4>
- White J. W., Hunt L.A. , Boote K. J., Jones J. W., Koo J., et al. **Integrated description of agricultural field experiments and production: The ICASA Version 2.0 data standards**. Computers and electronics in agriculture, 2013, vol. 96, p. 1-12.
- Pietragalla J., Valette L., Shrestha R., Laporte M.-A., Hazekamp T., Arnaud E. **Guidelines for creating crop-specific ontologies to annotate phenotypic data**, version 2.1, Alliance Bioversity International-CIAT, January 2022



# Questions de compétences

CQ Quels sont les objectifs d'un essai agronomique?

Par exemple, un essai concerne les tests de produits phytosanitaires pour lutter contre la maladie du black rot sur les ceps de vigne..

CQ Quels sont les facteurs étudiés?

Par exemple pour un essai agronomique de biocontrôle, les facteurs à étudier peuvent être le produit de biocontrôle, la date ou le stade de développement d'application du produit, la dose du produit à appliquer, .... Dans le cas du projet STAR, il s'agit d'essais agronomiques monofactoriels qui ne comportent donc qu'un seul facteur, le produit de biocontrôle à tester appliqué à la dose préconisée et à un stade fixé.

CQ Pour chaque facteur, quelles sont les modalités à tester?

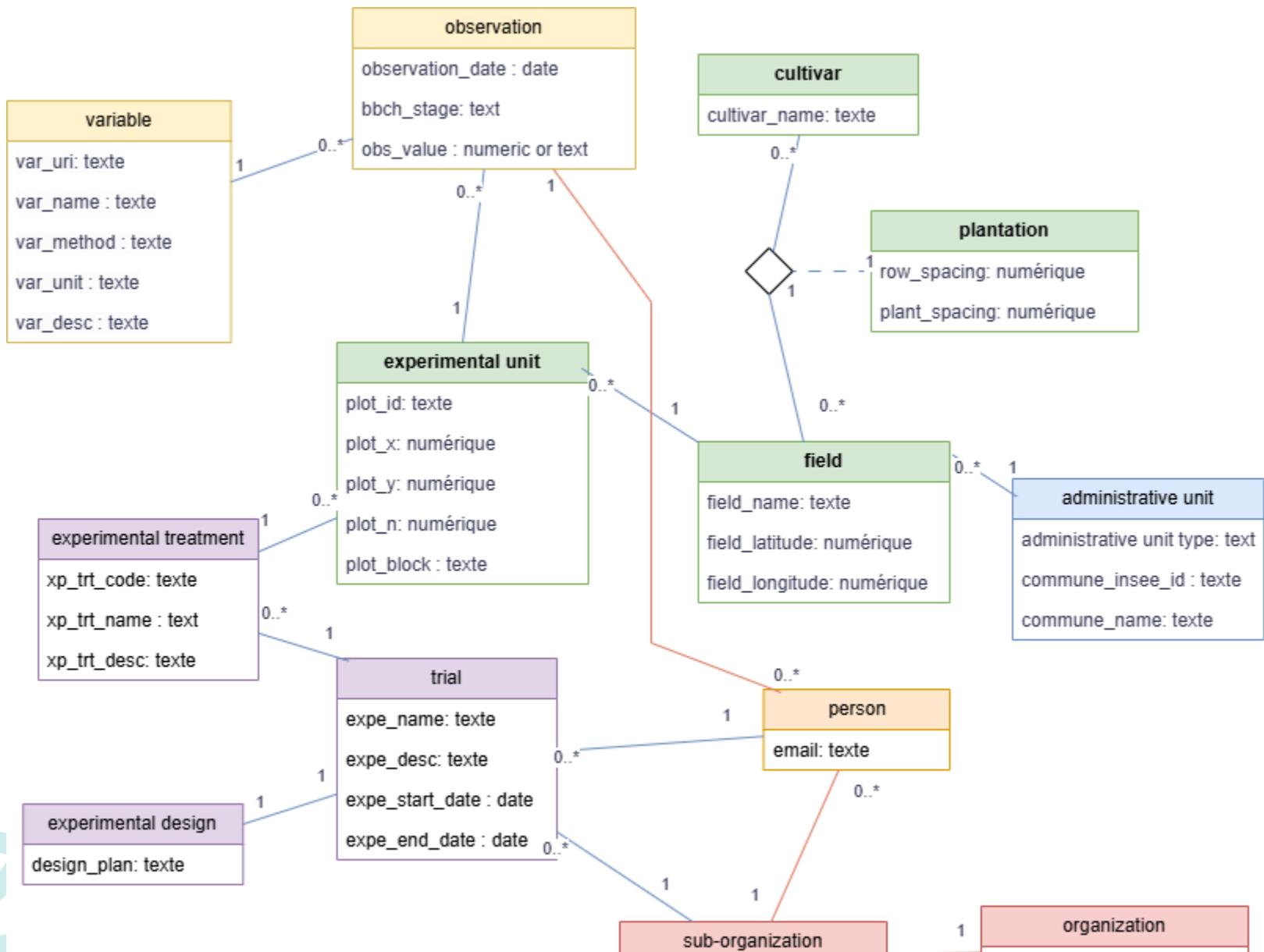
Par exemple, un essai en fertilisation azotée sur le blé a deux facteurs étudiés: la variété de blé (4 variétés) et la fumure (3 niveaux de fumure), donc il existe 4 modalités pour le premier facteur et 3 modalités pour le second facteur.

CQ Quels sont les différents traitements expérimentaux à mettre en oeuvre dans l'essai agronomique?

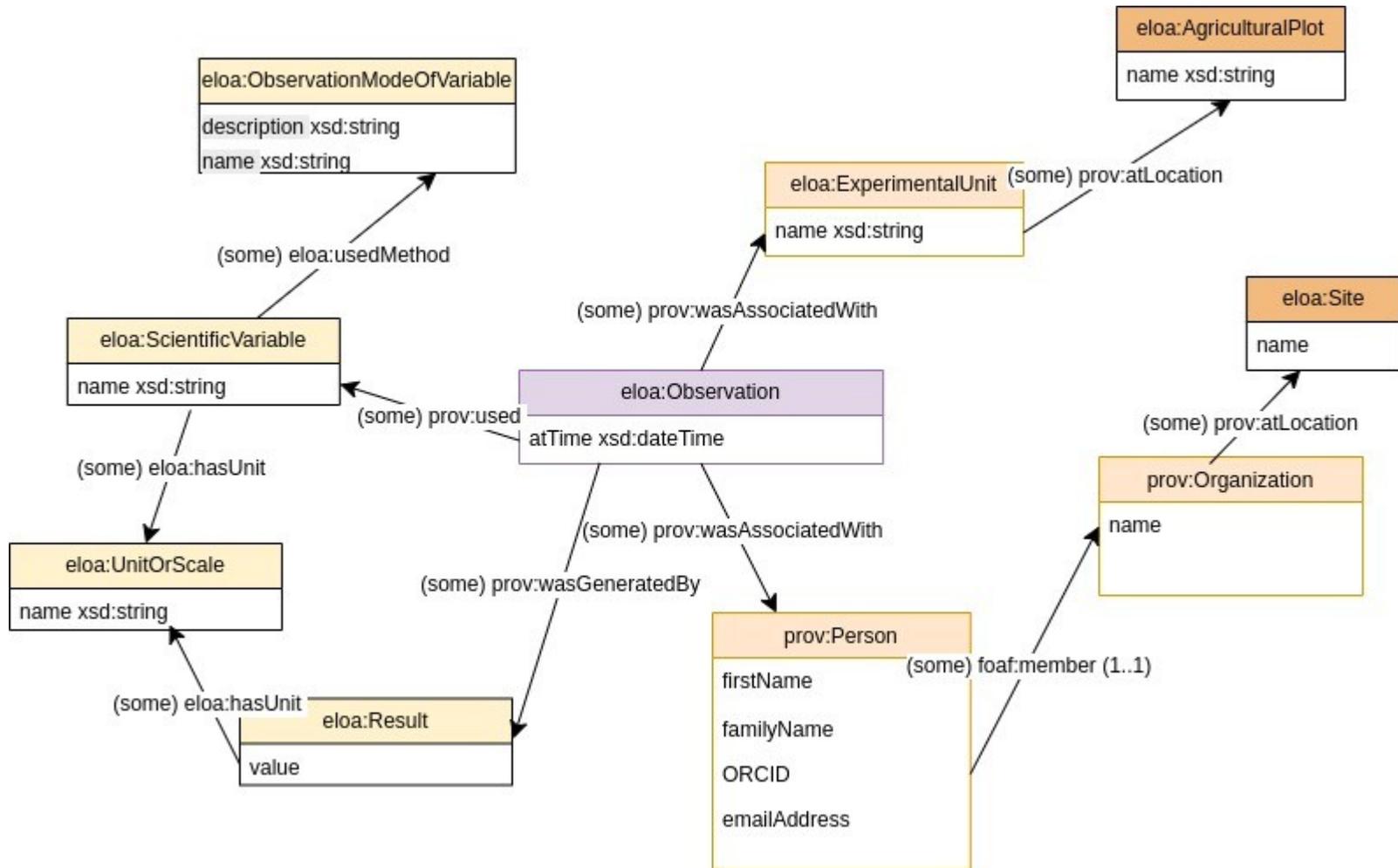
Dans l'exemple précédent, il y a donc 12 (4x3) traitements expérimentaux à mettre en oeuvre.



# Le modèle de données minimale



# Ontologie générique en expérimentation agronomique



# Action 2 :

## Outils pour la gestion des données

Modèles  
de saisies  
Excel

Fonction de génération de fichiers de saisie (publication open-source sur GitHub)  
Fichiers de saisie adaptés aux différents cas d'usage

Fonctions  
de qualité

Fonctions de qualité sous R (publication open-source sur GitHub)  
Enrichissement du package {qdd}  
Fonctions d'exploration graphique des données  
Données d'observatoires corrigées et validées

Applicatio  
n

- Application web ergonomique et intégrative des fonctions développées, en accès ouvert et gratuit

Outils  
d'interopé  
rabilité

Package R pour extraction et intégration dans R de jeux de données standardisés (publication open-source sur GitHub)  
Connecteur vers OpenSILEX  
• Connecteur On Farm Experiment (OFE)

# Action 3 :

## Accompagnement et transfert

### Formation

Guide d'aide en ligne et tutoriels vidéo pour la standardisation des données  
Fiches méthodo vigne et vin  
Guide d'aide à la décision pour le partage des données  
2 sessions de formations ouvertes et gratuites  
Formations internes IFV et Terres Inovia

### Publication de données

3 jeux de données standardisés et publiés en open-data  
Publication sur centre de ressources viti-oenologie  
Jeux de données références sur rd-agri

### Animation et communication

- Newsletter du projet
- Témoignages « success story » en lignes
- Séminaire ouvert et coorganisé avec OpenWineData
- Communications scientifiques et techniques en open-data

# Conclusion

- Tous: Modele de données d'experimentation: Travail en cours avec différents acteurs approches Bottom-Up et Top Down,
- Terres Inovia: alignement et publication des échelles phénologiques des oléagineux avec l'ontologie PPDO
- IFV, Elzeard, MISTEA: finir l'ontologie des bioagresseurs et produire le graphe des bioagresseurs de la vigne
- IFV, MISTEA: Décrire et nettoyer les variables phénologiques et les variables de stress biotique en Vigne
- Fichiers csv type avec Package R pour valider les données d'expérimentations en Biosolution
- Formations vis ma vie d'expérimentatrice.eur , de gestionnaire de données structurées
- Fiches méthodologiques sur l'observation de variables



# Remerciements



Florence AMARDEILH, Baptiste DARNALA



Llorenc CABRERA-BOSQUET, Pierre MARTRE



Xavier DELPUECH



Arnaud CHARLEROI, Pascal NEVEU Farzaneh KAZEMI-POUR, Catherine ROUSSEY, Anne TIREAU

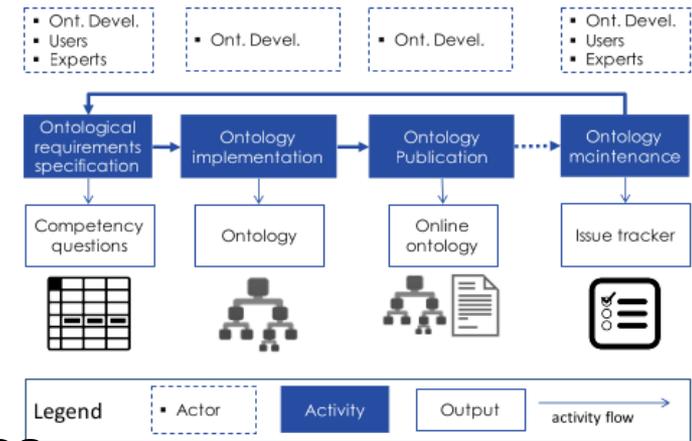


Frédéric SALVI, Martina AGUERA

# Méthode de conception ressources sémantiques

## Méthode « Linked Open Terms » (LOT)

- Questions de compétence (QC)
- Recherche de vocabulaires ou d'ontologies existantes
- **Recherche de définitions et de sources de référence**
- **Choix du niveau de formalisation OWL ↔ RDFS ↔ SKOS**
- Diagramme UML / CHOWLK
- Production de fichiers csv avec des experts +
- **Production de définitions qui explique le voisinage de l'entité dans l'ontologie**
- **Peuplement avec plugin Cellfie de Protege**
- Validation avec un raisonneur plugin SWRL Tab + drools
- **Validation avec SKOSPLAY**
- Publication : depot GIT + AgroPortal + Sparql EndPoint + DOI par RDG



Credit Espinoza-Arias, P., Poveda-Villalón, M. & Corcho, O. Using LOT methodology to develop a noise pollution ontology: a Spanish use case. J Ambient Intell Human Comput 11, 4557–4568 (2020).  
<https://doi.org/10.1007/s12652-019-01561-2>