



Contextualiser le vocabulaire

La sémantique de OWL est-elle suffisante ?

24 novembre 2017

Julie Bourbeillon
IRHS – Équipe Bioinformatique

■ L'Institut de Recherche en Horticulture et Semences

- TGU angevine – 14 équipes
- Pluridisciplinaire



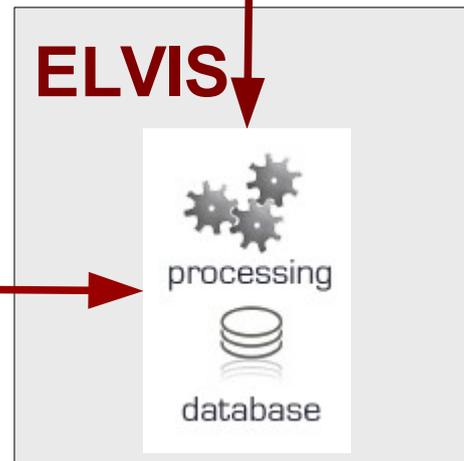
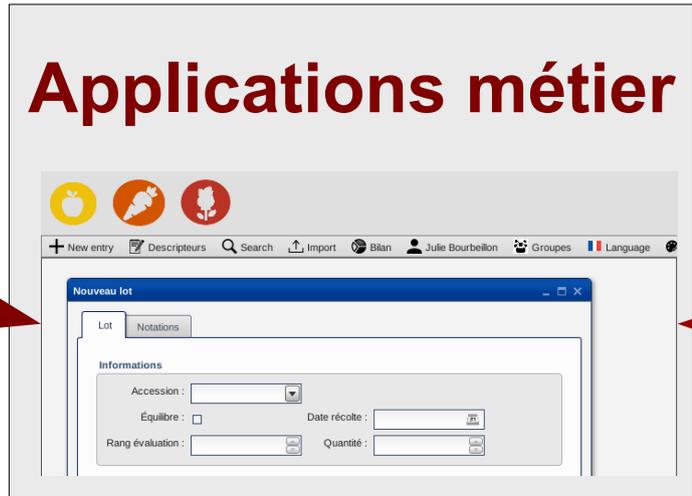
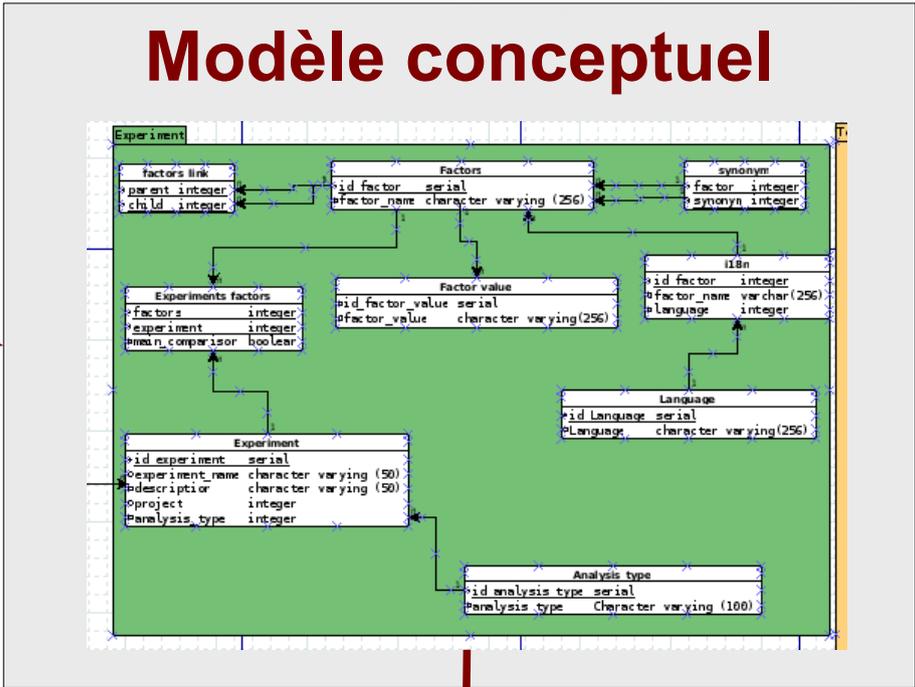
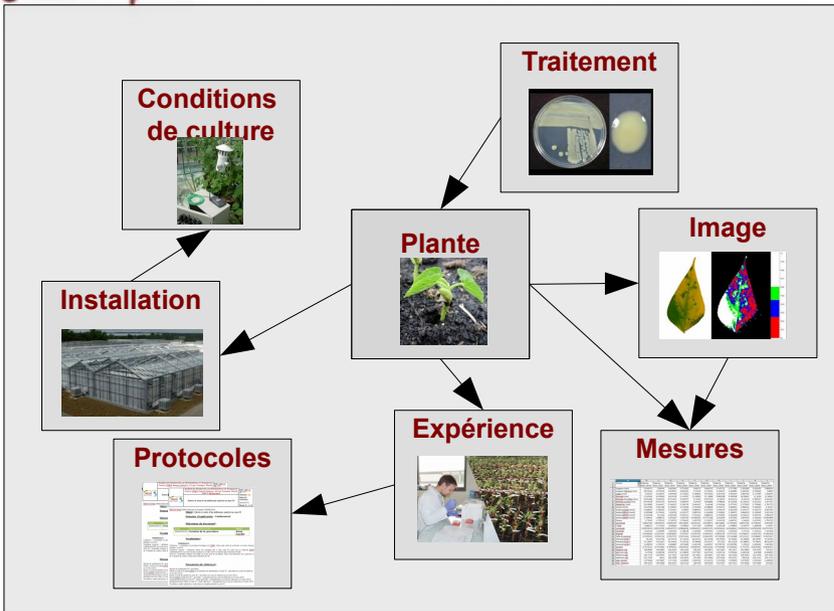
- Biologistes : généticiens, sélectionneurs, phytopathologistes, physiologistes, biochimistes, etc.
- Autres : modélisateurs, statisticiens, bioinformaticiens, physiciens, etc.
- De nombreux organismes étudiés



■ L'équipe bioinformatique

- Collaboration étroite avec les biologistes
- Développement de méthodes et outil pour la gestion et l'analyse de gros volumes de données complexes

ELVIS – Un LIMS pour l'IRHS





Limite → le vocabulaire

Description des échantillons

Prélèvement

Foliole

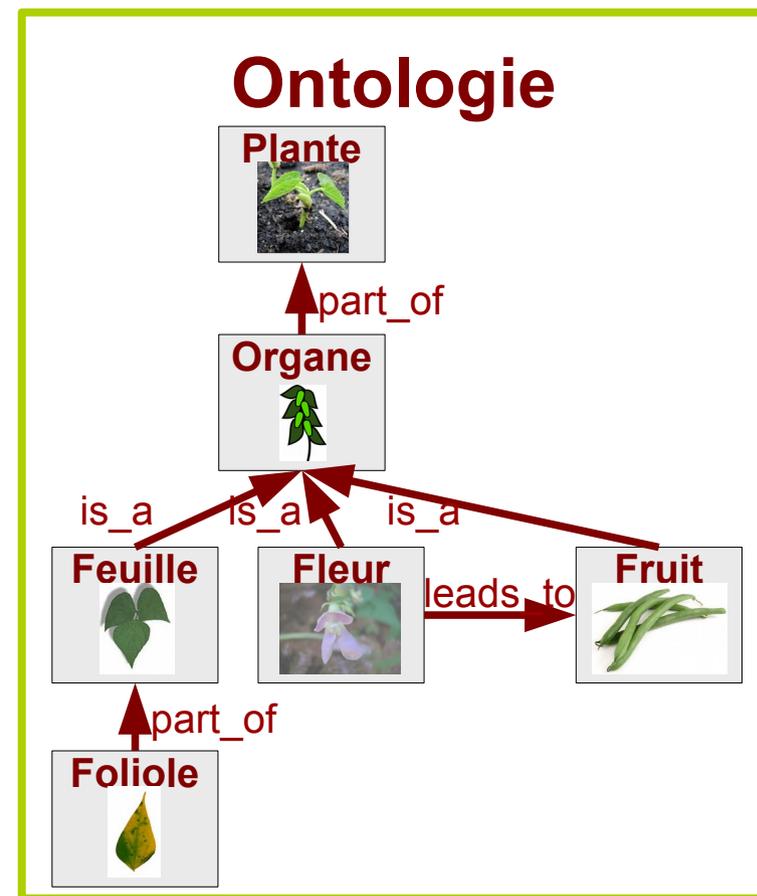
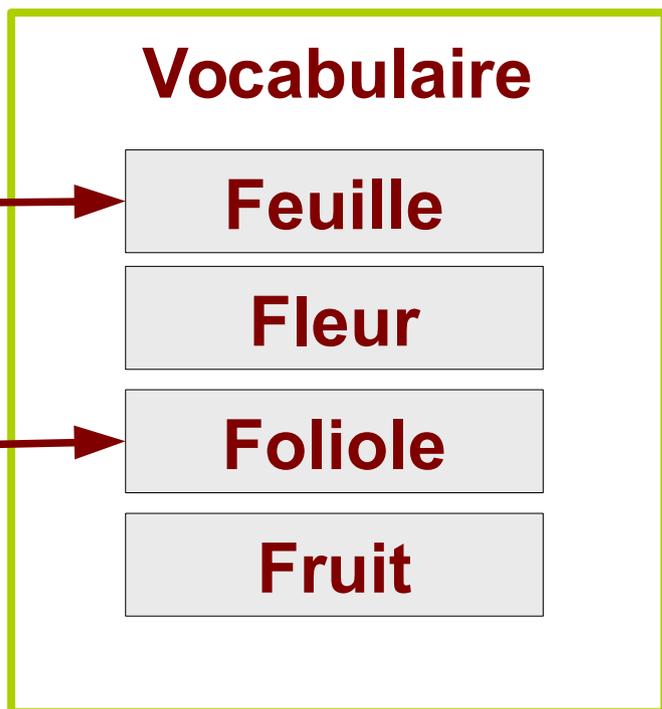
foliole

follole

Pièce foliaire

Retrouver toutes les images de foliole ?

Solution : ontologies



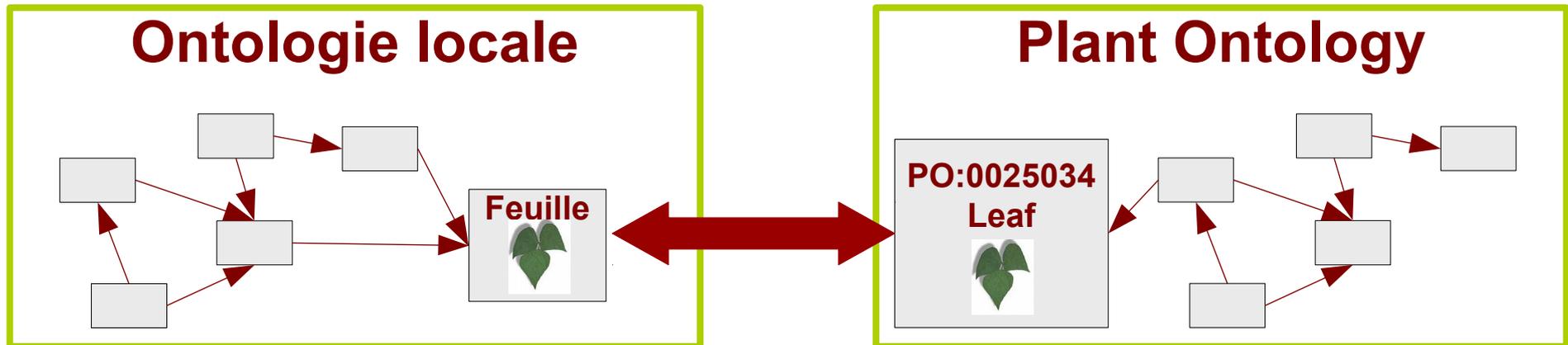


Ontologies utilisées

▪ Différents types d'ontologies

- Construction locale
 - Au plus près des besoins, spécifique
- Ontologies « de référence » → Efforts internationaux
 - Exemples : Gene Ontology, Plant Ontology, Crop Ontology, etc.
 - Sources : OBOFoundry, Bioportal (biomedical), Planteome (végétal)

▪ Liens entre les deux

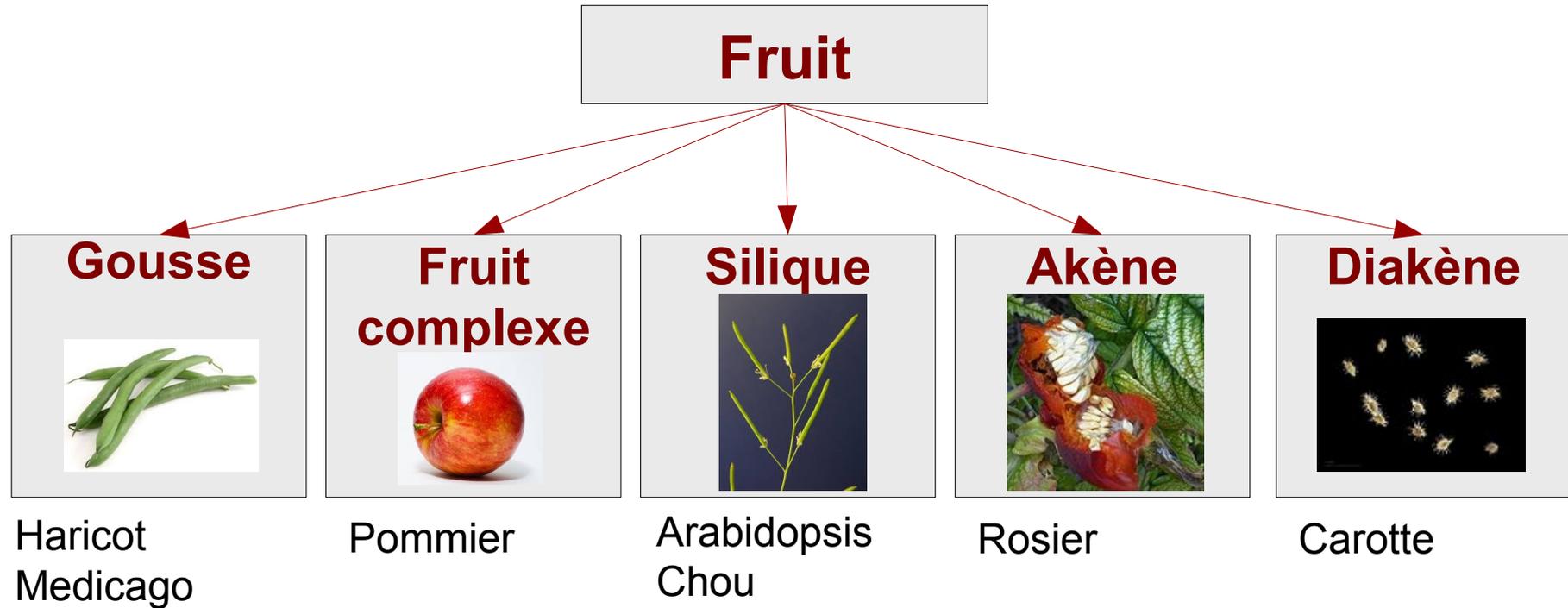


- Interopérabilité des ontologies
- Partage des données associées



Problème : flexibilité du vocabulaire « local »

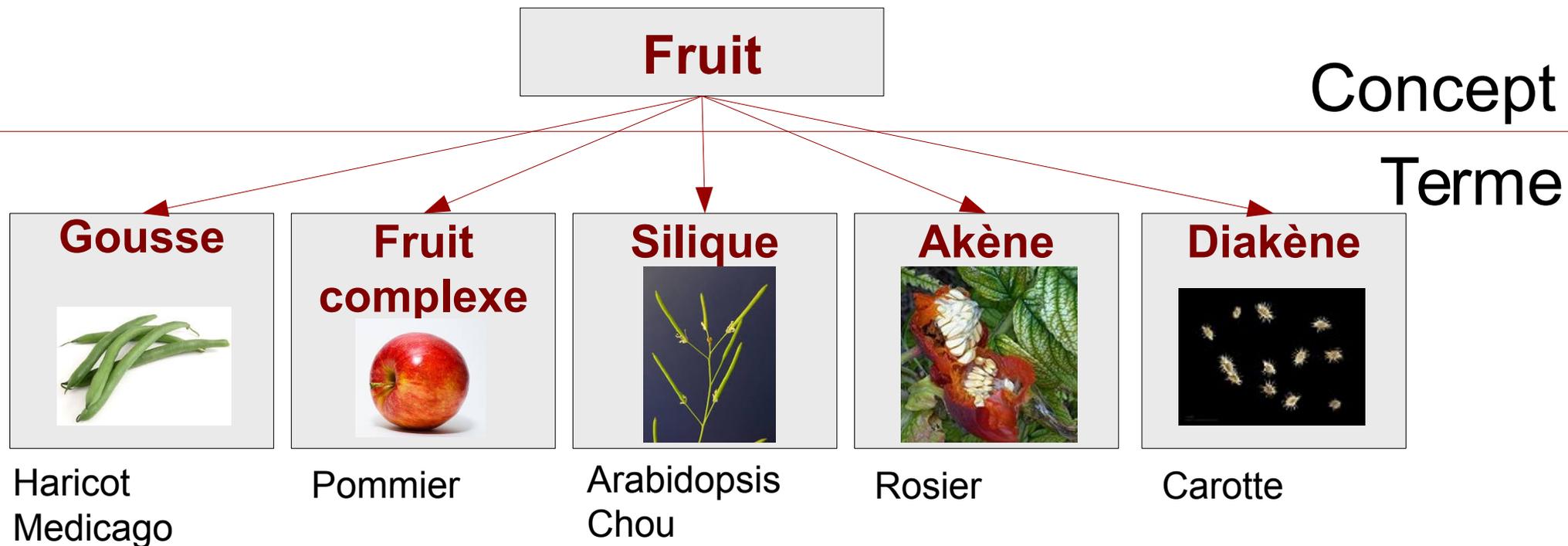
Flexibilité du vocabulaire : « Synonymes »



Conserver tous les mots associés au concept « Fruit » ?

Flexibilité du vocabulaire : « Synonymes »

Notion de terme

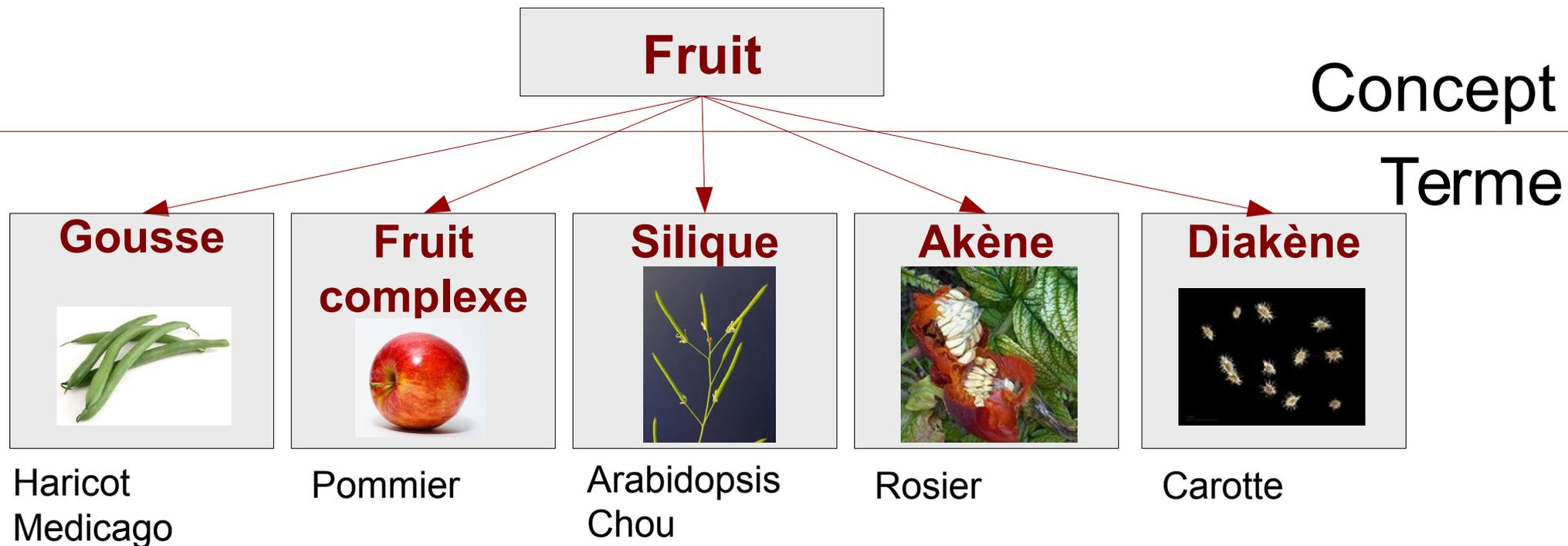


Un concept, plusieurs termes

- Conservation des synonymes
- Conservation des termes selon les cas
- Association des données au même concept, quel que soit l'organisme → études translationnelles

Flexibilité du vocabulaire : « Synonymes »

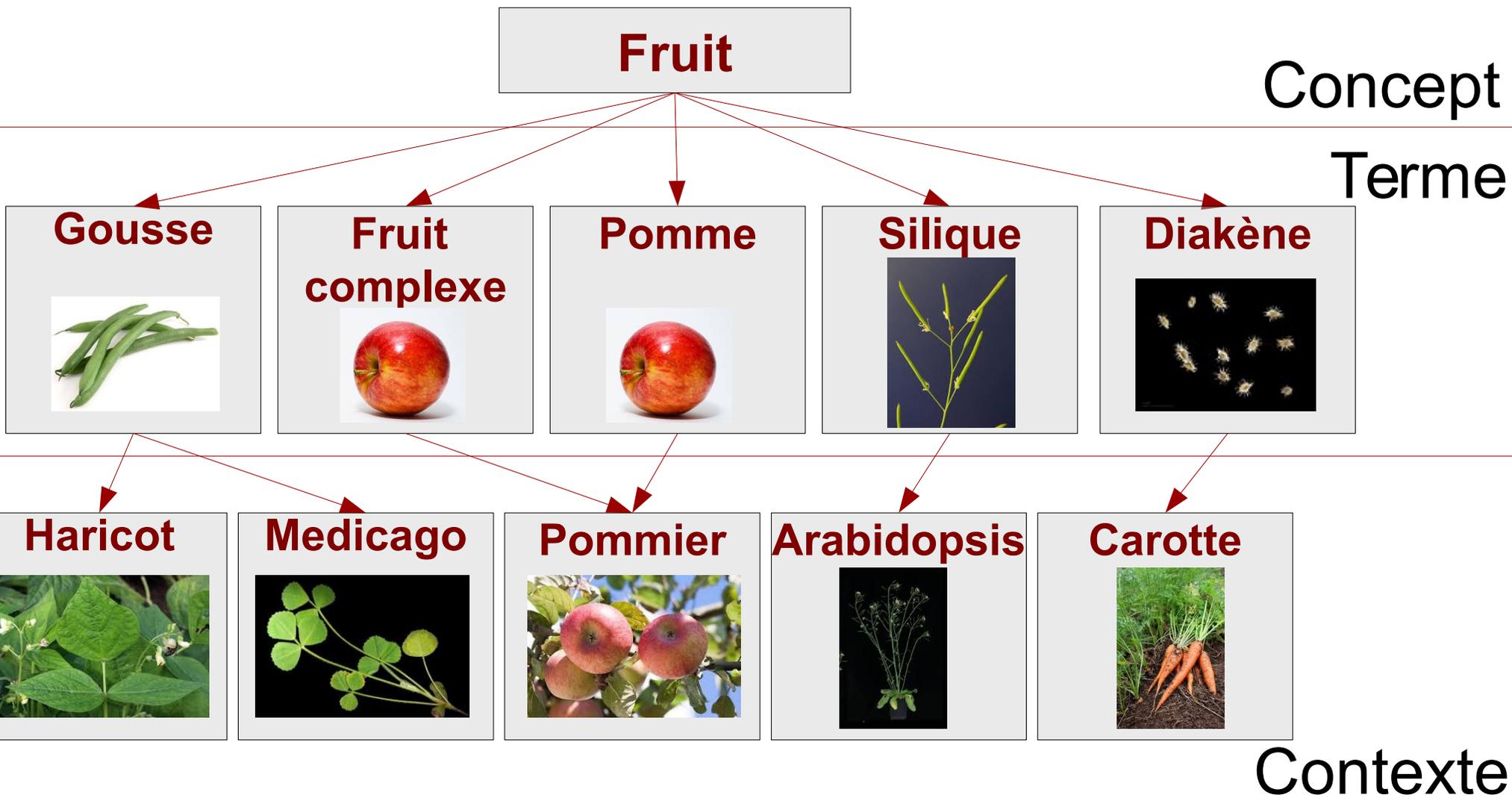
Spécificité des termes



Décrire à quel « cas » se rattache un terme ?

Flexibilité du vocabulaire : « Synonymes »

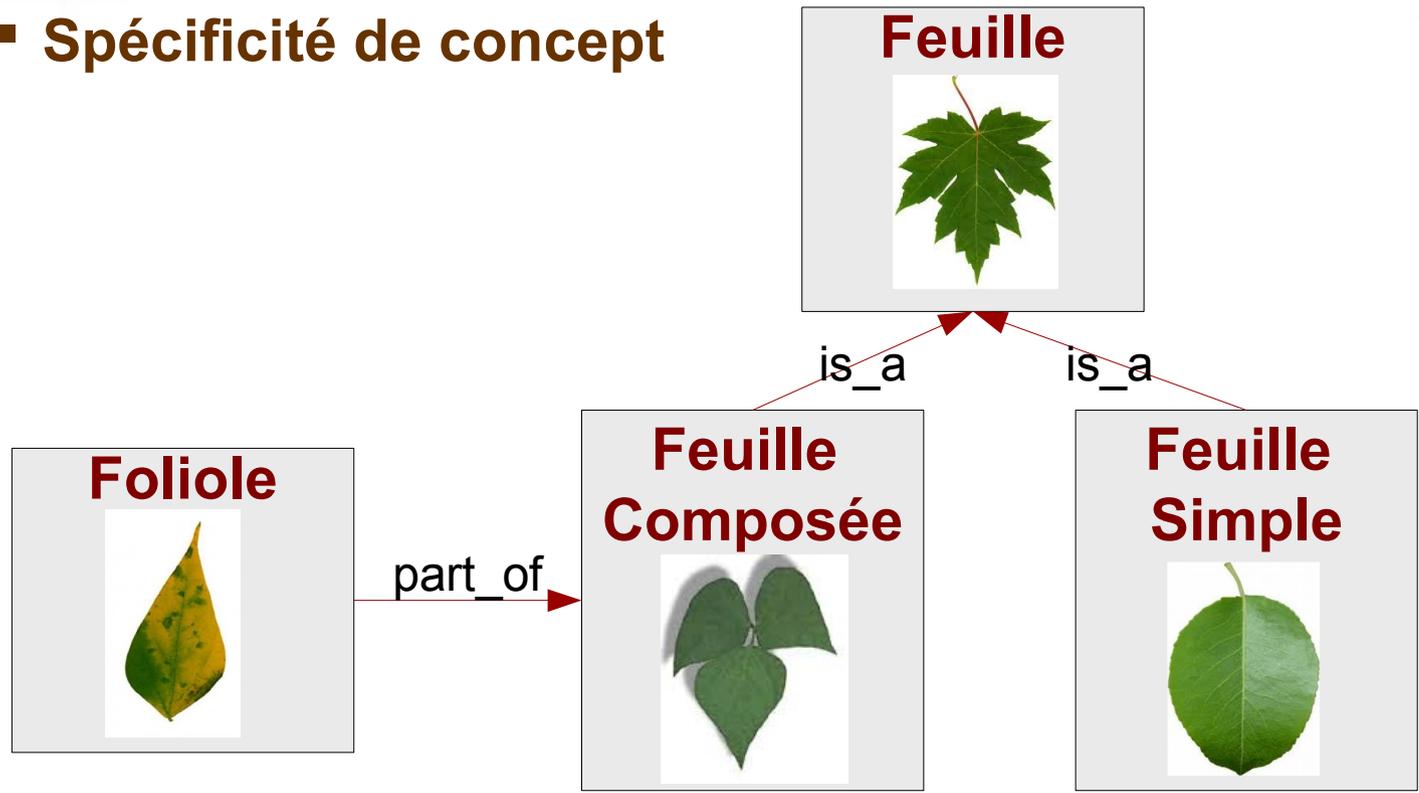
Notion de contexte





Flexibilité du vocabulaire : concepts spécifiques

▪ Spécificité de concept

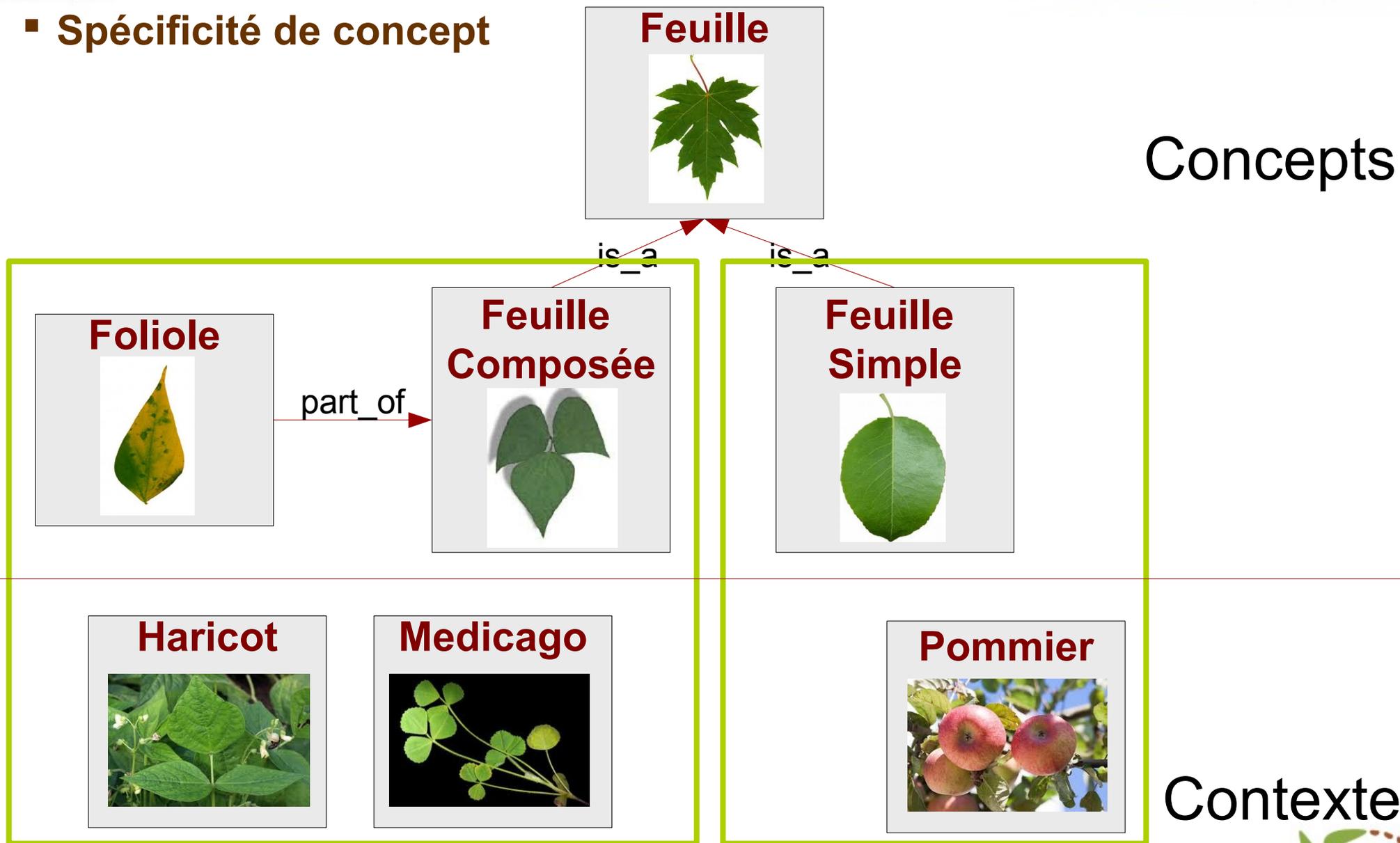


Concepts



Flexibilité du vocabulaire : concepts spécifiques

▪ Spécificité de concept



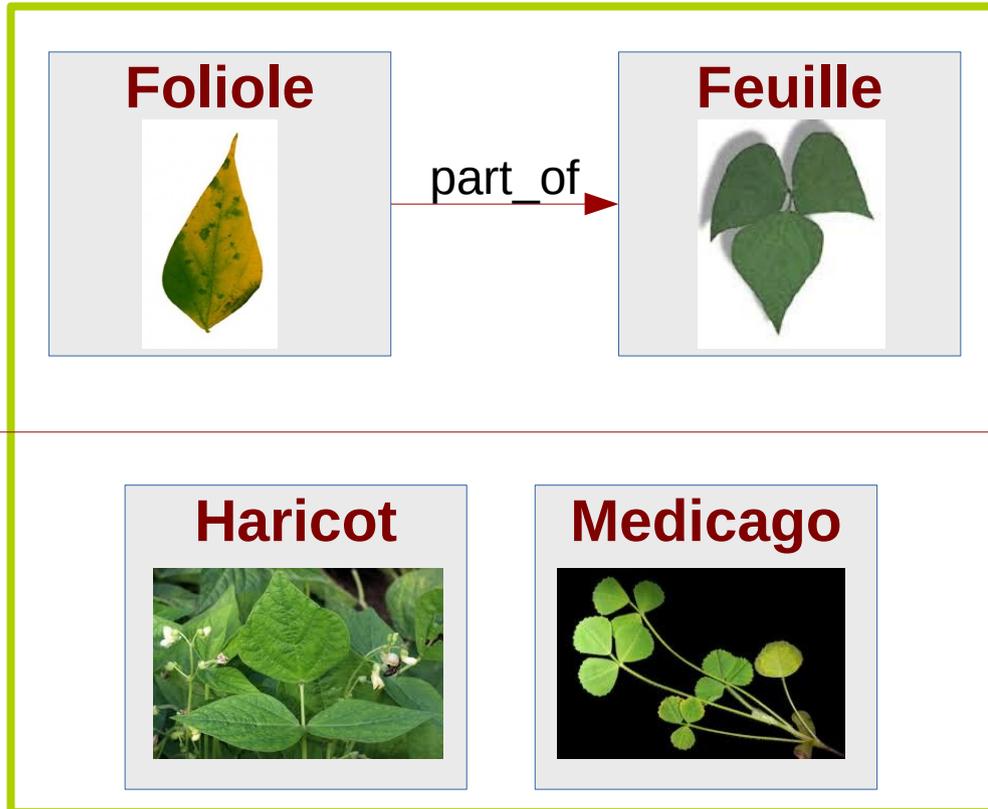
Contexte



Flexibilité du vocabulaire : concepts spécifiques

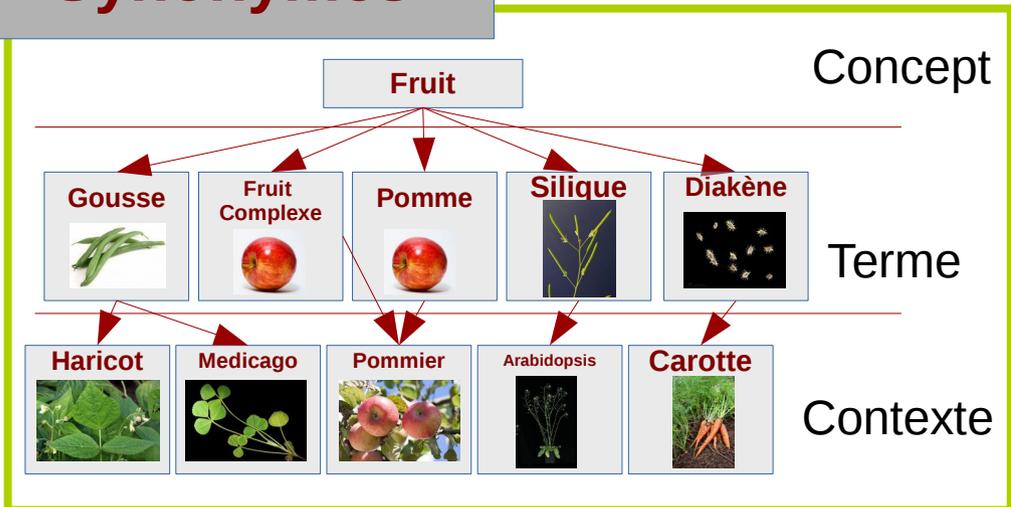
- **Spécificité de concept**
 - Usage « local »

Concepts

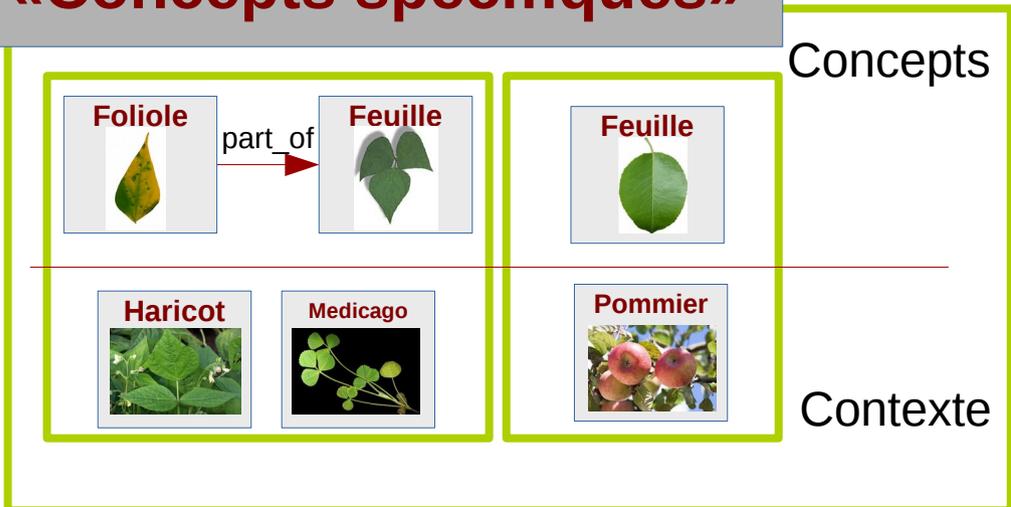


Contexte

« Synonymes »



« Concepts spécifiques »



Besoins

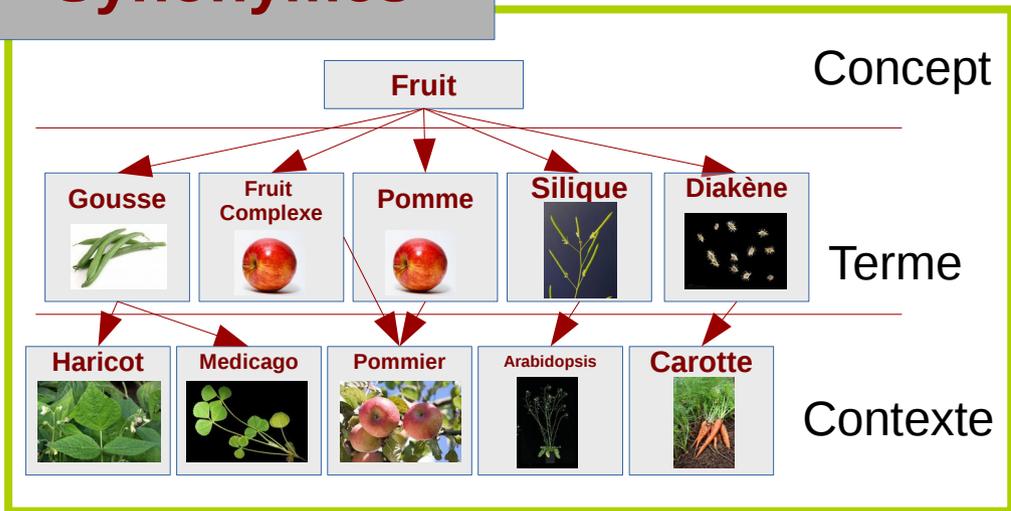
Saisie de données

- Présenter à l'utilisateur un vocabulaire adapté au contexte

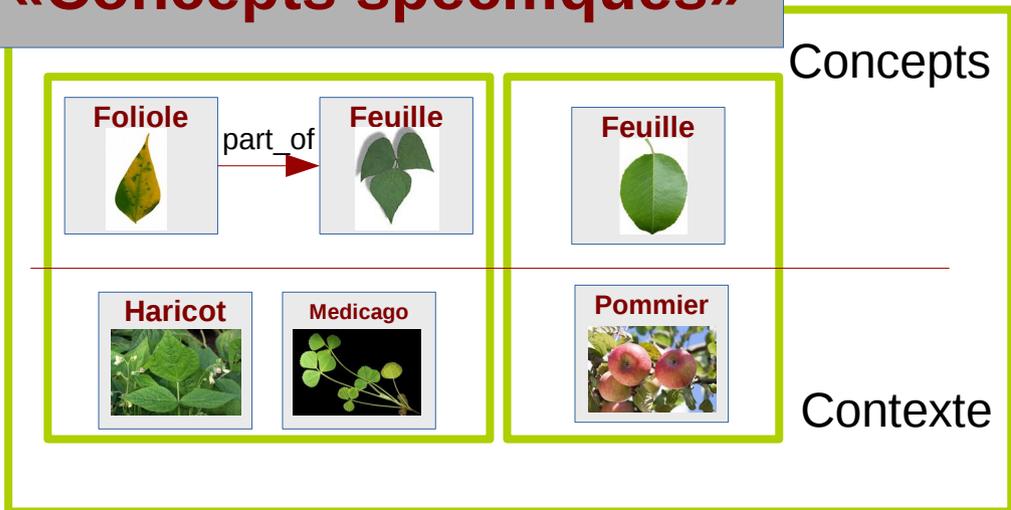
Analyse de données

- Raisonner sur les données à travers les contextes
- Permettre des études translationnelles

« Synonymes »



« Concepts spécifiques »



Représentation en OWL ?



Première approche

- **Modélisation « classique »**
 - Concept → Class
 - Term → Label

- **Ajout du contexte**
 - Nouvelle AnnotationProperty
 - pp:context

Première approche

Modélisation « classique »

« Synonymes »

Aj

The screenshot displays an ontology editor interface. On the left, a class hierarchy is shown with 'Fruit' selected. The hierarchy includes 'owl:Thing', 'Variable', 'Unit', 'Trial', 'Scientific_Object_Characteristics', 'Scientific Object', 'Object_Group', 'Object_Part', 'Plant_Part', 'Organ', 'Branch', 'Bud', 'Flower', 'Fruit', 'Leaf', 'Root', 'Organ_Part', 'Seed_Part', 'Seedling_Part', 'Primary Object', 'Person', 'Parameters', 'Location', 'Installation', 'Facility', 'Event', 'Equipment', and 'Document'. On the right, the 'Annotations: Fruit' panel shows three annotations:

- `rdfs:label` [language: fr] Fruit complexe
- `pp:context` [language: fr] Pommier
- `rdfs:label` [language: fr] Silique
- `pp:context` [language: fr] Arabidopsis
- `rdfs:label` [language: fr] Pomme
- `pp:context` [language: fr] Pommier
- `rdfs:label` [language: fr] Gousse
- `pp:context` [language: fr] Medicago



Première approche

▪ Modélisation « classique »

▪ Ajout

The screenshot displays a software interface for class modeling. On the left, a class hierarchy is shown under the heading 'Class hierarchy: Leaflet'. The hierarchy starts with 'owl:Thing' and branches into various categories like 'Document', 'Equipment', 'Event', etc. The 'Plant_Part' category is expanded, showing sub-classes like 'Organ', 'Organ_Part', 'Flower_Part', 'Fruit_Part', and 'Leaf_Part'. The 'Leaflet' class is highlighted in blue. On the right, the 'Class Annotations' panel for 'Leaflet' is visible, showing 'Annotations: Leaflet' and 'Annotations +'. Below this, the 'Description: Leaflet' section is shown, containing 'Equivalent To +', 'SubClass Of +', and 'Leaf_Part'.

« Concepts spécifiques »



Première approche

■ Modélisation « classique »

- Concept → Class
- Term → Label

■ Ajout du contexte

- Nouvelle AnnotationProperty
 - pp:context

■ Limites

- Annotations pas prises en compte par les raisonneurs
- Réintroduction du problème de vocabulaire
 - Pas de contrôle sur les valeurs de contexte
- Niveaux d'annotation différents pour les deux cas

The screenshot shows the Protege interface with the 'Fruit' class selected in the class hierarchy. The hierarchy includes 'owl:Thing', 'Variable', 'Unit', 'Trial', 'Scientific_Object_Characteristics', 'Scientific_Object', 'Object_Group', 'Object_Part', 'Plant_Part', 'Organ', 'Branch', 'Bud', 'Flower', 'Fruit', 'Leaf', 'Root', 'Organ_Part', 'Seed_Part', 'Seedling_Part', 'Primary Object', 'Person', 'Parameters', 'Location', 'Installation', 'Facility', 'Event', 'Equipment', and 'Document'. The 'Fruit' class is highlighted. The right pane shows annotations for 'Fruit', including 'rdfs:label' (Fruit complexe) and 'pp:context' (Pommier). Other classes like 'Silique', 'Pomme', and 'Gousse' are also annotated with 'pp:context' values.

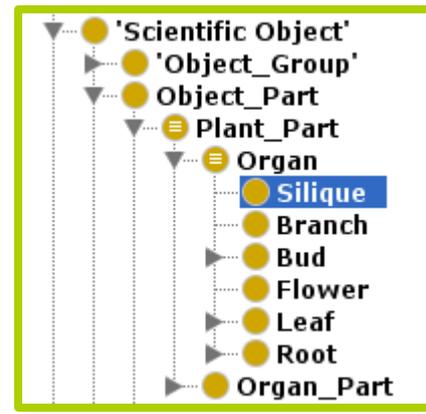
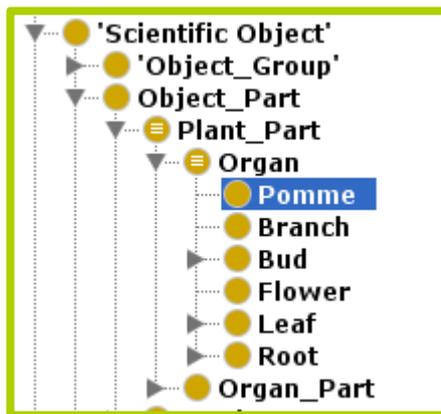
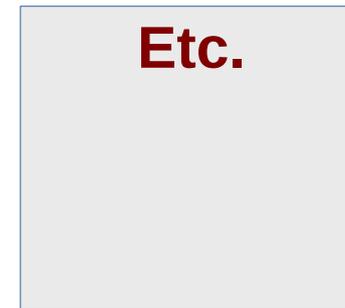
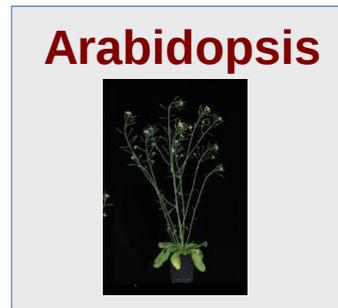
The screenshot shows the Protege interface with the 'Leaflet' class selected in the class hierarchy. The hierarchy includes 'owl:Thing', 'Document', 'Equipment', 'Event', 'Facility', 'Installation', 'Location', 'Parameters', 'Person', 'Scientific Object', 'Object_Group', 'Object_Part', 'Plant_Part', 'Organ', 'Branch', 'Bud', 'Flower', 'Fruit', 'Leaf', 'Root', 'Organ_Part', 'Flower_Part', 'Fruit_Part', 'Leaf_Part', 'Axil', 'Blade', 'Leaflet', 'Petiole', and 'Stipule'. The 'Leaflet' class is highlighted. The right pane shows annotations for 'Leaflet', including 'pp:context' (Arabidopsis) and 'Description: Leaflet'. It also shows 'Equivalent To' and 'SubClass Of' relationships.



Seconde approche

■ Une ontologie par contexte

- « synonyme » et « concept spécifique » → concept → class
- Equivalences de classes pour les correspondances



■ Limites

- Peu de différences
 - Grosse redondance
 - Enormément d'équivalences



Troisième approche

▪ Une ontologie « générique » + des spécificités

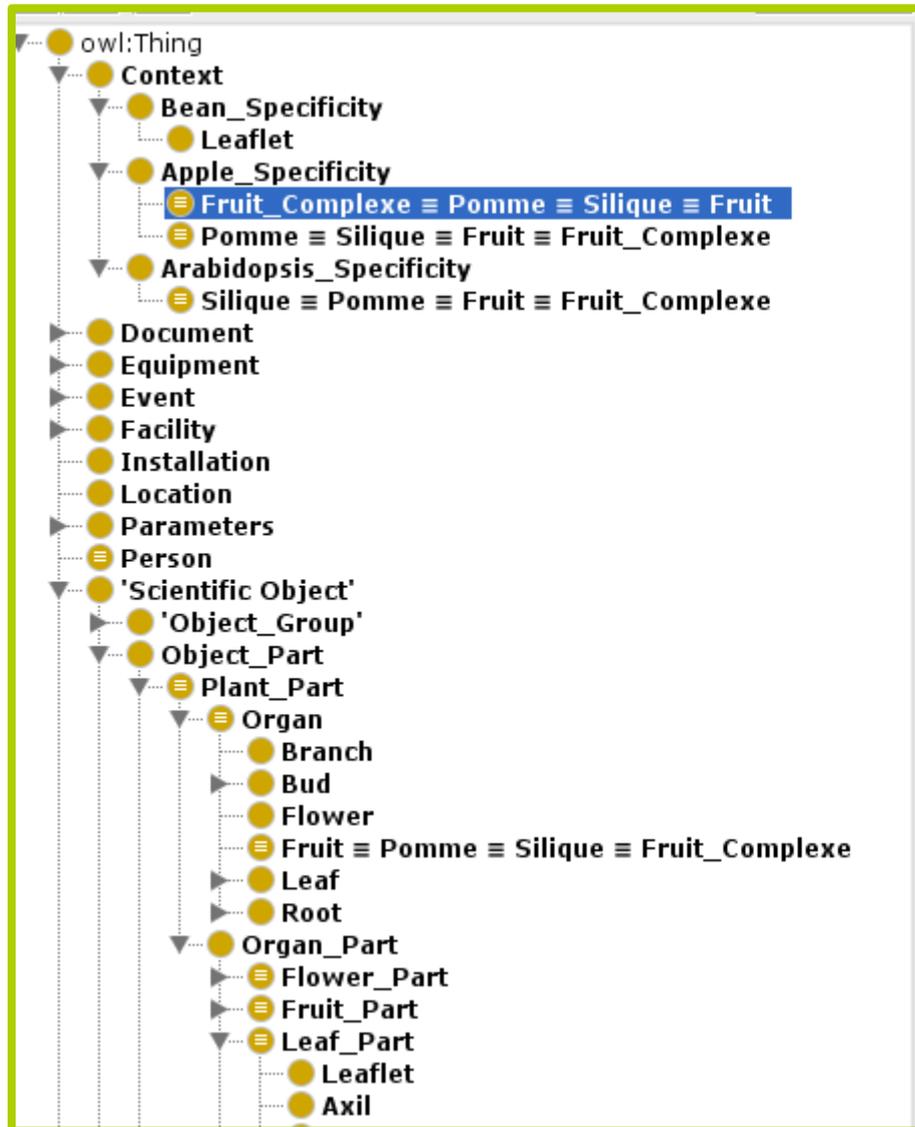
- « synonyme »
 - « synonyme » → concept → class dans une hiérarchie à part
 - Equivalence entre concept « local » et générique
- « concept spécifique »
 - Double relation is_a : hiérarchie générique + class une hiérarchie à part



Troisième approche

■ Une ontologie « générique » + des spécificités

« Synonymes »



Description: Fruit_Complexe

Equivalent To +
 Fruit

SubClass Of +
 Apple_Specificity

General class axioms +

SubClass Of (Anonymous Ancestor)

- Fruit
- Fruit
- part_of some Plant
- Silique
- Pomme
- Fruit_Complexe
- characterised_by some Organ_Development_Stage

Instances +

Target for Key +

Disjoint With +



Troisième approche

■ Une ontologie « générique » + des spécificités

« Concepts spécifiques »

The screenshot shows an ontology editor interface. On the left is a class hierarchy tree, and on the right is a detailed view of the 'Leaflet' class.

Class Hierarchy (Left):

- owl:Thing
 - Context
 - Bean_Specificity
 - Leaflet**
 - Apple_Specificity
 - Fruit_Complexe ≡ Pomme ≡ Silique ≡ Fruit
 - Pomme ≡ Silique ≡ Fruit ≡ Fruit_Complexe
 - Arabidopsis_Specificity
 - Silique ≡ Pomme ≡ Fruit ≡ Fruit_Complexe
 - Document
 - Equipment
 - Event
 - Facility
 - Installation
 - Location
 - Parameters
 - Person
 - 'Scientific Object'
 - 'Object_Group'
 - Object_Part
 - Plant_Part
 - Organ
 - Branch
 - Bud
 - Flower
 - Fruit ≡ Pomme ≡ Silique ≡ Fruit_Complexe
 - Leaf
 - Root
 - Organ_Part
 - Flower_Part
 - Fruit_Part
 - Leaf_Part
 - Leaflet
 - Axil

Class Detail View (Right):

Description: Leaflet

- Equivalent To +
- SubClass Of +
 - Bean_Specificity
 - Leaf_Part
- General class axioms +
- SubClass Of (Anonymous Ancestor)
 - part_of some Plant
 - part_of some Leaf
- Instances +
- Target for Key +
- Disjoint With +
- Disjoint Union Of +



Troisième approche

▪ Une ontologie « générique » + des spécificités

- « synonyme »
 - « synonyme » → concept → class dans une hiérarchie à part
 - Equivalence entre concept « local » et générique
- « concept spécifique »
 - Double relation is_a : hiérarchie générique + class une hiérarchie à part

▪ Limites

- Visibilité réduite sur les relations entre éléments spécifiques
- Equivalences croisées complexes et potentiellement intempestives
- Approche un peu différente dans les deux cas
- Difficulté à construire une ontologie spécifique à un contexte



Quatrième approche

- **Graphe de contextes + relation has_context**
 - Contexte → Class (Ensemble éventuellement hiérarchisé)
 - « synonyme »
 - Term → Concept enfant du concept « générique »
 - Relation has_context pour le lier au contexte
 - « concept spécifique »
 - Concept dans la hiérarchie “normale”
 - Relation has_context pour le lier au contexte



Quatrième approche

- **Graphe de contextes + relation has_context**
 - Contexte → Class (Ensemble éventuellement hiérarchisé)

The screenshot displays an ontology editor interface. On the left, a tree view shows a hierarchy of classes starting from 'owl:Thing'. The 'Pomme' class is highlighted in blue. On the right, the details for the 'Pomme' class are shown, including its description and various relationships.

owl:Thing

- Context
 - Vegetatively_Propagated_Plant
 - Rose_Bush
 - Apple
 - Seed_Propagated_Plants
 - Medicago
 - Arabidopsis
- Document
- Equipment
- Event
- Facility
- Installation
- Location
- Parameters
- Person
- 'Scientific Object'
 - 'Object_Group'
 - Object_Part
 - Plant_Part
 - Organ
 - Branch
 - Bud
 - Flower
 - Fruit
 - Gousse
 - Silique
 - Pomme

Description: Pomme

- Equivalent To +
 - has_context some Apple
- SubClass Of +
 - Fruit
- General class axioms +
- SubClass Of (Anonymous Ancestor)
 - part_of some Plant
 - characterised_by some Organ_Development_Stage
- Instances +
- Target for Key +
- Disjoint With +



Quatrième approche

▪ Graphe de contextes + relation has_context

- Contexte → Class (Ensemble éventuellement hiérarchisé)
- « synonyme »
 - Term → Concept enfant du concept « générique »
 - Relation has_context pour le lier au contexte
- « concept spécifique »
 - Concept dans la hiérarchie “normale”
 - Relation has_context pour le lier au contexte

▪ Limites

- Perte de l'équivalence entre concept générique et termes « locaux »
- Pas de différence entre une « vraie » relation de subsumption et nos spécialisations par contexte



Bilan

▪ **Réflexion actuelle**

- 4 approches possibles envisagées
- Aucune approche « idéale »

▪ **Questions**

- D'autres idées ?
- Une préférence ?

Merci pour votre attention et vos suggestions