



HAL
open science

OntoFruit : Ecophysiologie végétale de l'arboriculture fruitière – un référentiel documentaire indexé par une ontologie du domaine

Pascal Aventurier, Hugues H. Leiser, Herve Richard, Patrice Bellot

► **To cite this version:**

Pascal Aventurier, Hugues H. Leiser, Herve Richard, Patrice Bellot. OntoFruit : Ecophysiologie végétale de l'arboriculture fruitière – un référentiel documentaire indexé par une ontologie du domaine. Séminaire Texte et Connaissance, Jul 2008, Jouy-en-Josas, France. 32 diapositives. 16 p. hal-02822790


HAL Id: hal-02822790

<https://hal.inrae.fr/hal-02822790>

Submitted on 6 Jun 2020

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.




OntoFruit : Ecophysiologie végétale de l'arboriculture fruitière – un référentiel documentaire indexé par une ontologie du domaine

INRA Avignon:


- Pascal Aventurier (*Équipe Information Scientifique et Technique*)
- Hugues Leiser (*Unité Plante et Système Horticole*)
- Hervé Richard (*Unité Biostatistique et Processus spatiaux*)

Université Avignon:

- Patrice Bellot (*Laboratoire Informatique Appliquée*)



P Aventurier, H Leiser, H Richard.
Séminaire Texte et Connaissance à l'Inra – juillet 08





1

Plan

- ▶ Description du projet Ontofruit/Ontario :
 - Objectifs
 - Historique
- ▶ Ontologie :
 - Définitions
 - Quelques ontologies internationales
 - Construction de l'ontologie Ontofruit
- ▶ Le prototype (Ontario)
 - La recherche d'information
 - Le module d'annotation
- ▶ Conclusion

P Aventurier, H Leiser, H Richard.
Séminaire Texte et Connaissance à l'Inra – juillet 08



2

Objectifs (1/2)

Utiliser les technologies et outils du Web sémantique pour la représentation des connaissances d'un domaine scientifique (Ecophysiologie végétale appliquée à l'arboriculture fruitière)

Ceci dans le but de :

- ▶ Montrer que la recherche d'informations (RI) associée à une ontologie est plus performante qu'une recherche « classique ».
- ▶ Offrir aux chercheurs un outil visuel performant de gestion des connaissances scientifiques pour un domaine donné.



Objectifs (2/2)

Un outil de gestion des connaissances pour :

- ▶ Visualiser les relations sémantiques au sein d'une base de connaissances (BdC)
- ▶ Réaliser des recherches dans cette BdC
- ▶ Ajouter de nouvelles connaissances par l'annotation de nouveaux articles



Historique du projet (2006-2008)

- ▶ **2006**
 - Réalisation d'un Travail Pratique Encadré avec une étudiante (Inès Temou) de Master I IUP GMI – Traitement de l'Information Multimédia – Université d'Avignon
 - Faire un état de l'art des projets et des outils existants
 - Groupe de travail :
 - P. Aventurier,
 - H. Leiser (unité Plantes et Systèmes de culture Horticoles)
 - Hervé Richard (unité Biostatistique et Processus Spatiaux)
 - Patrice Bellot (Enseignant - Chercheur LIA Laboratoire en Informatique Appliquée de l'Univ. Avignon)
 - Inès Temou (stagiaire 2007)
- ▶ **2007**
 - Stage de 6 mois Master II – Inès Temou (financement 100% DV-IST)
 - Réalisation d'un 1er prototype pour la recherche d'information et les annotations, Construire une première ontologie (**Ontofruit**)
- ▶ **2008**
 - Stage de 6 mois Master II – Vivien Bernard (financement 100% DV-IST)
 - Amélioration du prototype



Ontofruit/Ontario

- ▶ Une étape « Recueil des connaissances » : **ONTOFRUIT**
Ontologie recueillant l'expertise autour du fruit
 - Construire une ontologie en collaboration avec les chercheurs de PSH (**Plante et Systèmes Horticoles**) puis actuellement du pôle PHI (**Pôle Horticole Intégré**)
- ▶ Une étape de développement informatique : **ONTARIO**
Outil de NavigaTion Automatique et de Recherche d'Information dans une Ontologie
 - Réaliser un module de **visualisation** et de **recherche** d'information de documents
 - Réalisation d'un module d'**annotation** des documents





Des ontologies pour décrire des connaissances quelques exemples

7

P Aventurier, H Leiser, H Richard.
Séminaire Texte et Connaissance à l'Inra – juillet 08



Ontologie (définition du point de vue informatique)

http://artist.inist.fr/article.php3?id_article=243

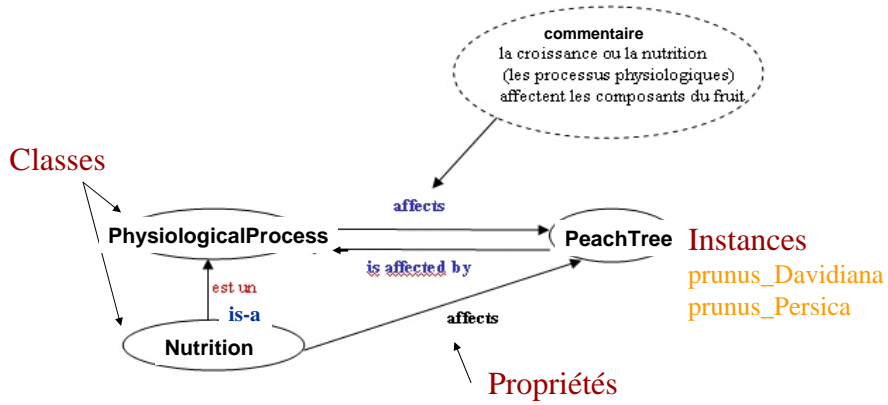
- **La spécification formelle et explicite d'une conceptualisation partagée**
- « **spécification explicite** », ensemble de concepts, de propriétés, d'axiomes, de fonctions et de contraintes explicitement définis.
- « **formelle** » précise que cette conceptualisation doit pouvoir être comprise et interprétée par un ordinateur.
- « **conceptualisation partagée** » précise l'aspect consensuel du vocabulaire employé
- « **conceptualisation** » représenter les connaissances, avec un objectif de réalisation.

8

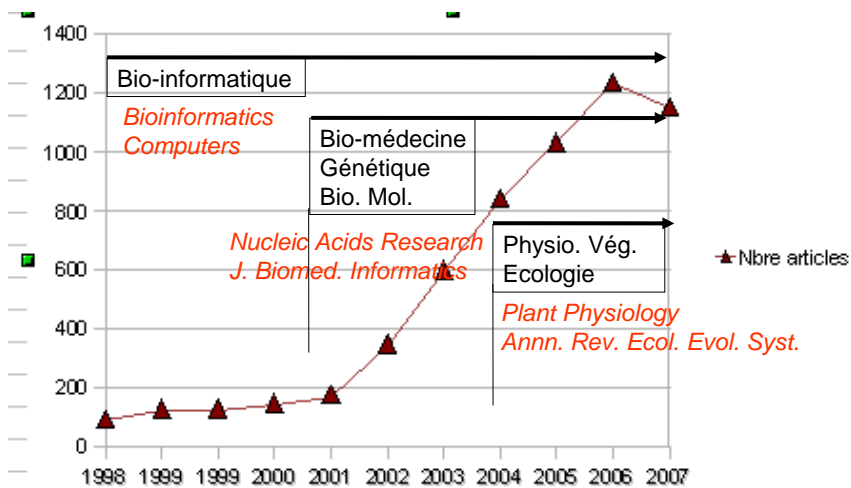
P Aventurier, H Leiser, H Richard.
Séminaire Texte et Connaissance à l'Inra – juillet 08




Ontologie (exemple)



Web sémantique et Ontologie : évolutions du nombre des publications (source WOS : 1998-2007)



Les Projets existants

- ▶ **The Gene Ontology Project** <http://www.geneontology.org/>
 - Regroupe 3 bases relatives aux données recueillies pour des organismes modèles (drosophile, souris, levure saccharomyces)
 - Identifie génétiquement 3 types de d'activités: fonctions moléculaires, processus biologiques fondamentaux, compartiments cellulaires de localisation
- ▶ **Plant Ontology** <http://www.plantontology.org>
 - Vocabulaire contrôlé structuré en ontologie dans le domaine de la biologie végétale au sens large afin d'exploiter le contenu des bases de données dispersées dans ce domaine. Utilise les méthodes de Gene Ontology Consortium.
- ▶ **Unified Medical Language Systems (UMLS)** <http://semanticnetwork.nlm.nih.gov/>
 - Développe des systèmes d'information favorisant l'intégration sémantique de langages hétérogènes en biomédecine et santé.
 - Produit différents vocabulaires et lexiques ainsi que des outils pour les exploiter.
- ▶ **Open Biomedical Ontology (OBO)** : <http://www.obo.sourceforge.net>.
 - Définit un cadre générique qui permet à d'autres projets de s'insérer dans une organisation collective. Plant Ontology et Gene Ontology font partie de ce projet.



Ontofruit: construction de l'ontologie

- 1) Le référentiel terminologique du domaine et la structure conceptuelle construits à partir:
 - ▶ d'ontologies existantes : *Plant Ontology*; *Gene Ontology*
 - ▶ de connaissances « académiques » (ouvrages, traités)... enrichis avec:
 - ▶ les instances et concepts extraits du corpus d'articles
- 2) La construction des relations suit la même démarche que pour les concepts (existant + analyse corpus)
 - ▶ L'objectif est de produire un nombre limité de relations, mais pertinentes, notamment pour la caractérisation des processusexemple: influences / interacts / produces



Ontofruit: annotation de l'ontologie

1) Construction du fichier d'annotation de départ :

- ▶ Les instances (mots-clés) candidates: lecture de l'article
- ▶ Les relations: lier les entités dans Ontofruit

A	B	C	D	E	F
* observation	<i>instance</i>	concept	<i>[property]</i>	<i>instance</i>	concept
	<i>agriCultural_Practice</i>	Cultural_Practice	<i>influences</i>	<i>Physiological_Process</i>	Plant
	<i>arthropod</i>	biotic_factor	<i>interacts_with</i>	<i>pear</i>	Plant
	<i>branch</i>	Aerial_System	<i>influences</i>	<i>aerial_system</i>	Aerial_System
	<i>cacopsylla_pyri</i>	biotic_factor	<i>interacts_with</i>	<i>pear</i>	Plant
	<i>composition_of_population : richness</i>	Ecology	<i>characterizes</i>	<i>arthropod : cacopsylla_pyri</i>	Biotic_Factor
	<i>distance_from_the_edge</i>	biotic_factor	<i>interacts_with</i>	<i>population</i>	Ecology
	<i>explanatory_variables</i>	<i>method_techinc</i>	<i>analyzes</i>	<i>population</i>	Ecology
	<i>hedge</i>	Biotic_Factor	<i>interacts_with</i>	<i>pear</i>	Plant
	<i>management_practices</i>	Cultural_Practice	<i>influences</i>	<i>Physiological_Process</i>	Plant
	<i>multivariate_analysis</i>	<i>Method_Technics</i>	<i>analyzes</i>	<i>population : population_varia</i>	Ecology


2) Conversion des données dans Ontofruit:

- ▶ Traitement du fichier par un programme: vérification de l'intégrité (a) des relations instances/concepts posés et (b) des relations (propriétés)





Le prototype Ontario



P Aventurier, H Leiser, H Richard.
Séminaire Texte et Connaissance à l'Inra – juillet 08

INRA

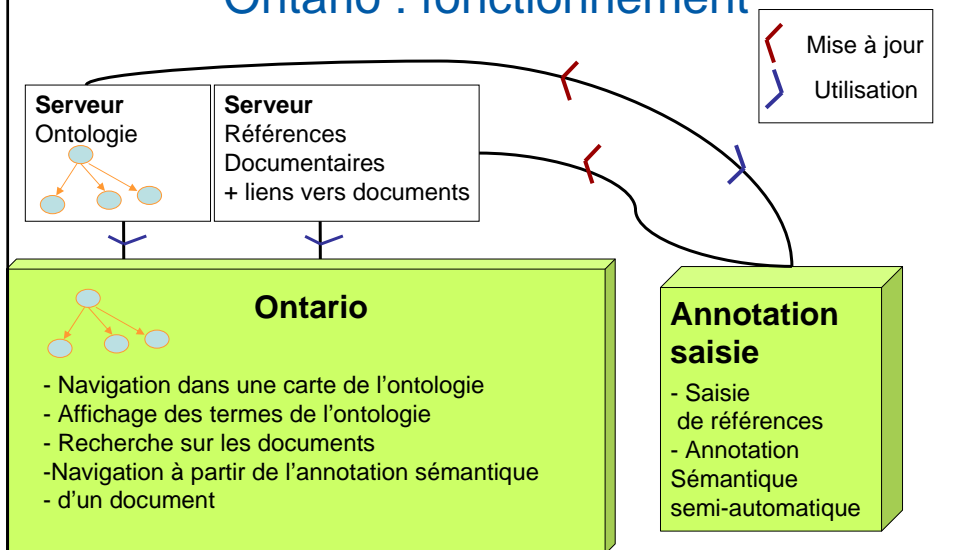
18

Le prototype d'Ontario: fonctionnalités

- ▶ L'utilisateur peut
 - Rechercher dans les concepts et visualiser la partie de l'ontologie correspondante puis accéder aux documents référencés
 - Afficher les différents concepts qui décrivent chaque document
 - Annoter et relier un nouveau document à l'ontologie (enrichissement de **Ontofruit**)
- ▶ Les concepts sont liés aux documents
 - Métadonnées dans la BDD locale
 - Textes en ressources externes (via un URI)



Ontario : fonctionnement



Affichage complète

Affichage de l'annotation du document

Affichage du document annoté avec l'ontologie

Visualisation graphique de l'ontologie

Le Module d'annotation

- ▶ Permet d'ajouter un document dans la base
- ▶ Identifie les concepts présents dans le document en vert (moteur lucene)
- ▶ Propose à l'utilisateur de valider la relation entre 2 concepts si les deux sont présents dans l'ontologie
- ▶ Propose à l'utilisateur de décrire lui-même des relations entre au moins un « concept candidat » présent dans le texte mais pas dans l'ontologie
- ▶ Met à jour la base de documents et l'ontologie



Module Annotation (2)

Plant and Soil (2006) 282:117–126
DOI 10.1007/s11104-005-5373-7

© Springer 2006

Spatio-temporal variations in axial conductance of primary and first-order lateral roots of a maize crop as predicted by a model of the hydraulic architecture of root systems

A. Pierret^{1,4}, C. Doussan² & L. Pagès³

¹UR Solutions, IRD-IWMI-NAFRI, BP 06, Vientiane, LAO PDR. ²Soil and Environment, INRA Climate, Domaine St Paul, Site Agroparc, 84914, Avignon cedex 9, FRANCE. ³INRA Plants and cultural Systems in Horticulture, Site Agroparc, 84914, Avignon cedex 9, FRANCE. ⁴Corresponding author*

Received 10 August 2005. Accepted in revised form 21 November 2005

Key words: 3D modelling, hydraulic conductance, root architecture, water uptake, xylem

Abstract

Rates at which water can be transported along plant roots (axial pathway) vary through time, in part depending on xylem maturation. Because of experimental constraints, the dynamics of root functional heterogeneity under field conditions remains mostly uncharted territory. Recent advances in mechanistic modelling offer opportunities to bypass such experimental limitations. This paper examines the dynamics of local variations in axial conductance of primary and first-order lateral roots of a maize crop using the architecture-based modelling approach developed by Doussan et al. (*Annals of Botany*: 81, 213–223, 1998). Specifically, we hypothesised that points of major resistance to long distance water transfers could arise from discrepancies between the hydraulic maturity (or water carrying capacity) of main axes and branch roots. To test this assumption, spatial distributions of root axial conductance were tested after 30, 60 and 100 days at soil depths of 10, 50 and 100 cm under a maize (*Zea mays* L.) crop sown at a density of 8 plants m⁻². As the crop developed, the corresponding root populations encompassed ever increasing amounts of hydraulically mature first-order laterals (branch roots): after a 100-day growth period, the vast majority of laterals had reached their maximum axial conductance at all soil depths down to 100 cm. In contrast, the axial conductance of a large proportion of main axes (primary roots) remained low, even at shallow soil depths and after 100 days of growth. The imbalance between the hydraulic maturity of primary and lateral roots was most conspicuous at soil depths of 100 cm, where ~10% only of the former compared to ~80% of the latter, had reached their maximum axial conductance after a 100-day growth period.

P Aventurier, H Leiser, H Richard.
Séminaire Texte et Connaissance à l'Inra – juillet 08



24

Module Annotation (3)

Annotario 0.2

Remplissez les différents champs, en séparant les différents mots clefs et auteurs par des ';' ou des ','.

Metadonnees	abstract
URL :	<input type="text" value="http://dx.doi.org/10.1007/s11104-005-5373-7"/>
titre :	<input type="text" value="Primary and First-order Lateral Roots of a Maize Crop as Predicted by a Model of the Hydraulic Architecture of Root Systems"/>
Auteurs :	<input type="text" value="Pierret, A.; Doussan, C.; Pagès, L."/>
Annee :	<input type="text" value="2006"/>
Source :	<input type="text" value="Plant and Soil"/>
Mots clefs :	<input type="text" value="3D modelling - hydraulic conductance - root architecture - water uptake - xylem"/>
Type :	<input type="text" value="Journal article"/>
Base :	<input type="text" value="Springer"/>

P Aventurier, H Leiser, H Richard.
Séminaire Texte et Connaissance à l'Inra – juillet 08



25

Module Annotation (4) : Résultats de l'analyse

Relations proposées

Étape suivante pour construire une relation ontologique non présente dans l'analyse automatique

- periodic_growth (Growth) affects root (Root_System)
- periodic_growth (Growth) affects lateral_root (Root_System)
- aerial_system (Aerial_System) isInfluencedBy branch (Branch)
- hydraulic_architecture (Root_System) interactsWith water_uptake (Absorption)
- hydraulic_architecture (Root_System) interactsWith water (Abiotic_Factor)
- root (Root_System) isAffectedBy periodic_growth (Growth)
- root (Root_System) isAffectedBy shoot_growth (Growth)
- malate (Organic_Acid) isInfluencedBy potassium (Component)
- temporal_variation (Method_Technic) analyzes hydraulic_architecture (Root_System)
- temporal_variation (Method_Technic) analyzes water_uptake (Absorption)
- temporal_variation (Method_Technic) analyzes axial_conductance (Transport)
- crop_protection (Cultural_Practice) interactsWith crop (Cultural_Type)
- shoot_growth (Growth) affects root (Root_System)
- shoot_growth (Growth) affects lateral_root (Root_System)

Conclusion

Bilan

- ▶ Démarche privilégiant la « finesse » de l'annotation sémantique
- ▶ Bien adaptée aux ontologies de domaine pour des corpus relativement faibles
- ▶ L'application Ontario est générique et peut travailler sur n'importe quelle ontologie

Perspectives

- ▶ Evaluer la réelle contribution des utilisateurs experts pour faire évoluer le couple connaissances ontologiques / documents avec Ontario
- ▶ A confronter avec d'autres expériences
- ▶ Construction en cours de l'ontologie du Pôle Thématique d'Avignon PHI (Horticulture intégrée)

