



HAL
open science

Application des principes FAIR à l'aide à la sélection d'un emballage alimentaire : un exemple de démarche Open Science à l'Inra

Patrice Buche, Julien Cufi, Stéphane Dervaux, Juliette Dibie, Valérie
Guillard, Luc Menut, Bruno Perret, Michard Rakotoson

► To cite this version:

Patrice Buche, Julien Cufi, Stéphane Dervaux, Juliette Dibie, Valérie Guillard, et al.. Application des principes FAIR à l'aide à la sélection d'un emballage alimentaire : un exemple de démarche Open Science à l'Inra. Cahier des Techniques de l'INRA, 2018, N° Spécial: Données de la recherche, 13 p. hal-04693770

HAL Id: hal-04693770

<https://hal.inrae.fr/hal-04693770v1>

Submitted on 10 Sep 2024

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.



Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License

Application des principes FAIR à l'aide à la sélection d'un emballage alimentaire : un exemple de démarche Open Science à l'Inra

Patrice Buche¹, Julien Cufi¹, Stéphane Dervaux², Juliette Dibie²,
Valérie Guillard¹, Luc Menut¹, Bruno Perret³, Michard Rakotoson³

Résumé. Nous présentons dans cet article la mise en œuvre d'une démarche Open Science qui permet de collecter dans un entrepôt de données ouvert sur le Web des données provenant de sources hétérogènes en suivant les principes FAIR de gestion des données (i.e. Findable, Accessible, Interoperable, Reusable). L'interopérabilité et la réutilisation des données est garantie par l'utilisation d'une ontologie pour leur annotation dans un vocabulaire de référence et leur structuration en tableaux. L'association d'un DOI (Digital Object Identifier) permet de rendre accessibles les tableaux de données annotés par l'ontologie, en particulier par l'intermédiaire du moteur de recherche Datacite. Nous illustrons notre démarche Open Science par la collecte de données caractérisant des matériaux d'emballage commercialisés ou innovants, leur réutilisation dans un outil d'aide à la sélection multicritères d'emballages alimentaires de produits frais respirant (e.g. abricot, fromage, endives) et leur accès dans des data paper.

Mots clés : Open science, ontologie, aide à la décision, data paper, emballages alimentaires

Introduction

La démarche Open Science proposée est guidée dans cet article par une question sociétale émanant des acteurs de la filière des emballages alimentaires. Nous proposons de répondre à la question suivante : « Je veux un emballage alimentaire qui préserve mon produit, fait à partir de ressources renouvelables, sans OGM, si possible transparent avec un coût inférieur à 3 € / kg ». Ce travail a été réalisé à la demande de l'équipe ePOP (Eco-efficient Polymeric & Organic Packaging) de l'UMR IATE qui met au point de nouveaux emballages plus écologiques issus de ressources renouvelables non alimentaires et biodégradables en conditions naturelles. Notre objectif est de mettre à disposition des solutions informatiques adaptées pour permettre aux acteurs de la filière des emballages alimentaires (e.g. scientifiques concepteurs d'emballages innovants, centres techniques, fabricants d'emballages, industriels de l'agro-alimentaire) une comparaison facilitée de matériaux d'emballages et une réduction du nombre d'expérimentations à réaliser. Notre deuxième objectif est d'augmenter la visibilité du travail des chercheurs en rendant accessible par des moteurs de recherche (par exemple Datacite) les données qu'ils souhaitent mettre à disposition et par la publication de data papers en complément des articles habituels. Dans cet article, nous présentons tout d'abord notre démarche de collection de données provenant de sources hétérogènes dans un entrepôt de données ouvert sur le Web guidée par une ontologie avec le logiciel @Web. Puis, nous présentons un exemple de réutilisation des données annotées avec @Web dans un outil d'aide à la décision multicritères permettant de sélectionner des emballages alimentaires pertinents pour un aliment donné. Enfin, nous présentons

¹ UMR Ingénierie des Agropolymères et Technologie Émergente, CIRAD, Inra, Montpellier SupAgro, Université Montpellier, 34060 Montpellier, France ; patrice.buche@inra.fr, julien.cufi@inra.fr, guillard@univ-montp2.fr, luc.menut@inra.fr

² UMR Mathématique et Informatique Appliquées, AgroParisTech, Inra, 75231 Paris, France ; stephane.dervaux@inra.fr, juliette.dibie@agroparistech.fr

³ UMR Génie et Microbiologie des Procédés Alimentaires, Inra, AgroparisTech, 78850 Thiverval-Grignon, France ; bruno.perret@inra.fr, michard.rakotoson@inra.fr

les fonctionnalités développées dans l'outil @Web pour associer un identifiant pérenne aux données, les rendre accessibles par des moteurs de recherche et faciliter leur publication dans des data papers.

@Web, un outil d'annotation de données provenant de sources hétérogènes

@Web est un outil logiciel Web de collection et d'annotation de données guidée par une ontologie dont l'objectif est de structurer des données de même type provenant de sources hétérogènes dans des tableaux afin de faciliter leur intégration, leur interopérabilité et leur réutilisation (Buche et al., 2013). La **Figure 3** illustre ce principe : des caractéristiques d'emballages provenant de fiches fournisseurs (voir exemple **Figure 1**) et d'articles scientifiques (voir exemple **Figure 2**) sont collectés et annotés dans un format tabulaire standardisé et uniforme à l'aide d'une ontologie de domaine.



Technical data sheet

Product: DanaMAP 04-01

Composition of Materials

Material	Thickness μm	Unit Weight g/m^2
OFA	15	17,3
Adhesive	1,7	2
PE	60	55
Total (Rounded)	-	74



Typical values

Test	Standard	Unit	Value
Lamination strength	ASTM F 904	N/15 mm	> 1,5
Tensile strength	ASTM B62	N/15 mm	MD 45-55 / CD 40-45
Puncture resistance	ASTM D 1306 Mod. to Danapak version	N	18-20
Water vapour transmission rate	ASTM F 1249 38°C, 90% RH	$\text{g/m}^2/24\text{h}$	8
CO ₂ transmission rate	ASTM D 3985 23°C, 0% RH	$\text{cm}^3/\text{m}^2/24\text{h}/\text{bar}$	210
Oxygen transmission rate	ASTM D 3985 23°C, 0% RH	$\text{cm}^3/\text{m}^2/24\text{h}/\text{bar}$	50-60

Figure 1. Caractéristiques d'un emballage commercial (danapak), format PDF, anglais.

Table 3. Main characteristics multilayered biobased films.

Film	Shape (tr = transparent) n-tr = non transparent)	Permeability		Thickness ^a (µm)
		O ₂ ^b (cc/m ² d) 23°C – 75% RH	H ₂ O ^b (g/m ² d) 38°C – 90% RH	
Natureflex™ N913 (cellulose-based)	Flex. film (tr)	9.9	10.1	55
Natureflex™ N931 (cellulose-based)	Flex. film (n-tr)	3.4	5.0	44
Ecoflex + Ecovio/ Ecovio/Ecoflex + Ecovio	Flex. film (n-tr)	815.0	216.4	55
Metallized PLA	Flex. film (n-tr)	25.4	2.3	20
Cellophane™/ Metal layer/PLA	Flex. film (n-tr)	9.1	9.7	46
Paper/AlOx/PLA	Flex. film (n-tr)	45.7	6.0	91
Bioska 504 (multilayer PLA)	Flex. film (tr)	617.6	275.1	34
Natureflex™/PLA	Flex. film (tr)	11.01	11.3	60
Cellophane™/PLA	Flex. film (tr)	10.5	13.8	100
PHB/Ecoflex	Flex. film (tr)	142.1	80.6	87
Xylophane A (coated on paper)	Flex. film (n-tr)	3.7	24.3	100 (coating = 9)
Xylophane B (coated on paper)	Flex. film (n-tr)	6.0	23.9	100 (coating = 9)
PLA tray	Tray (tr)	46.8	3.8	200–300

Figure 2. Caractéristiques d'emballages innovants (Peelman et al, 2013), format HTML, anglais.

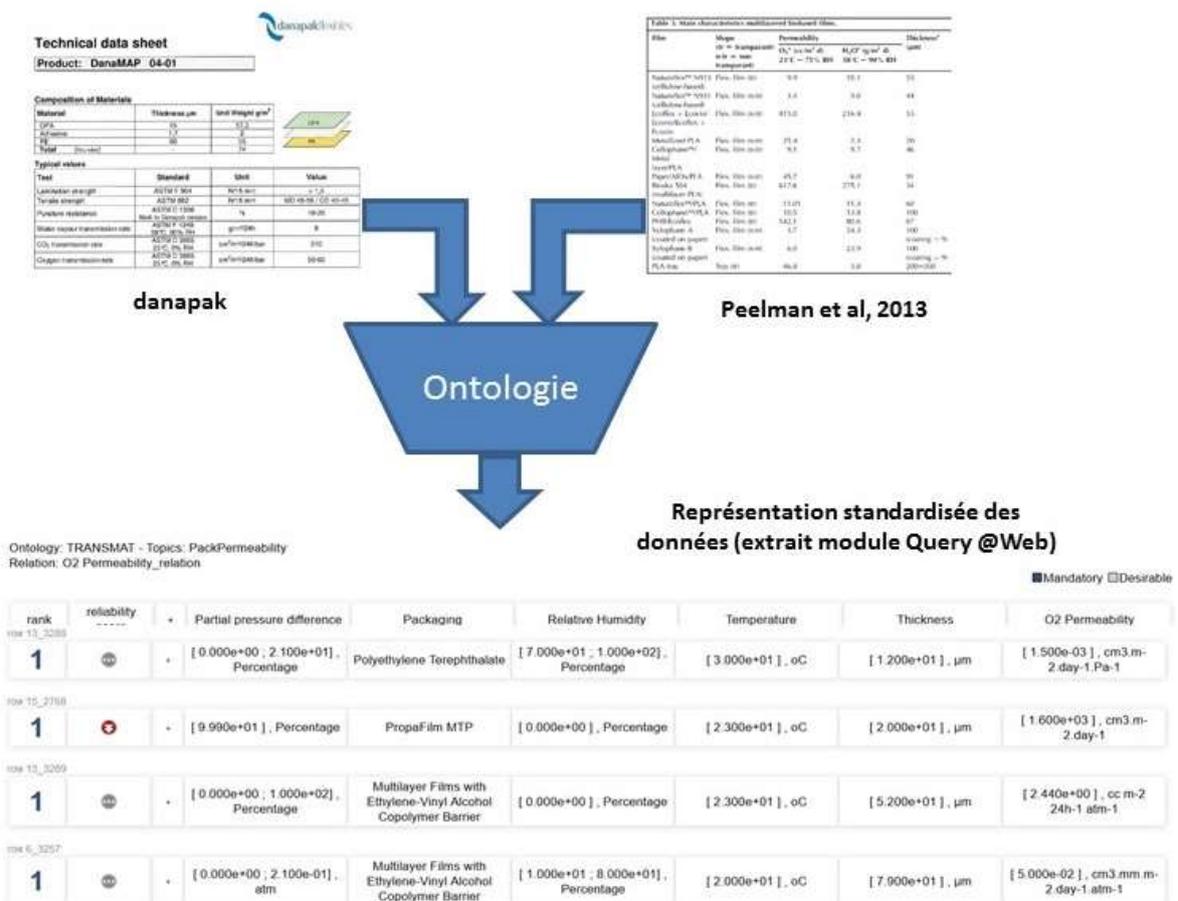


Figure 3. Intégration de sources de données hétérogènes guidée par ontologie.

Dans @web, l'annotation et l'interrogation des données sont réalisées grâce à une ressource termino-ontologique (RTO). Une RTO est composée d'une ontologie à laquelle est associée une composante terminologique multilingue. Dans @Web, la RTO permet de définir comment structurer les données en tableau et le vocabulaire de référence utilisé pour les annoter. La structure d'un tableau est définie à partir d'une ou plusieurs relation(s) n-aire(s) pré-définies dans la RTO. Les colonnes du tableau correspondent aux différents arguments de la ou des relation(s). Dans @Web, nous nous intéressons aux données quantitatives expérimentales dans le domaine des sciences du vivant, ce qui requière en particulier la définition dans la RTO de quantités avec leurs unités de mesure. Par exemple, dans la **Figure 4**, l'extrait de tableau annoté dans @Web contient plusieurs mesures de perméabilité à l'oxygène de matériaux d'emballage. Chaque colonne correspond à un argument de la relation n-aire *O2PermeabilityRelation* définie dans l'ontologie. Dans la quatrième colonne du tableau se trouve la valeur de la perméabilité à l'oxygène de l'emballage défini dans la première colonne dans des conditions expérimentales données par la différence de pression partielle à l'oxygène et son épaisseur (respectivement troisième et cinquième colonnes).

n°	Packaging	Method	Partial pressure difference Unit : Pa	O2 Permeability Unit : cc m-2 24h-1	Thickness Unit : m
1	Cellulose	O2 permeability method	[0.000e+00 ; 1.013e+05]	[1.700e+01 ; 1.850e+01]	[2.000e-05 ; 2.200e-05]
2	Cellulose	O2 permeability method	[0.000e+00 ; 1.013e+05]	1.700e+01	[2.900e-05 ; 3.100e-05]
3	Polyester	O2 permeability method	[0.000e+00 ; 1.013e+05]	[5.000e+01 ; 1.300e+02]	2.500e-05
4	Polyester	O2 permeability method	[0.000e+00 ; 1.013e+05]	[9.000e+00 ; 1.500e+01]	2.500e-05
5	Polyolefin	O2 permeability method	[0.000e+00 ; 1.013e+05]	7.800e+03	2.500e-05
6	Polyolefin	O2 permeability method	[0.000e+00 ; 1.013e+05]	2.600e+03	2.500e-05
7	Cellophane Film	O2 permeability method	[0.000e+00 ; 1.013e+05]	3.000e+00	2.100e-05

Figure 4. Extrait de tableau annoté dans @Web.

La relation n-aire *O2PermeabilityRelation* utilisée pour annoter le tableau de la **Figure 4** peut être représentée par le schéma de la **Figure 5**. C'est l'une des 57 relations n-aires de la RTO de domaine *Matter Transfer*⁴. Cette relation n-aire permet de représenter la perméabilité à l'oxygène d'un emballage dans des conditions expérimentales données par son épaisseur, son humidité relative, la différence de pression partielle à l'oxygène et la température ambiante.

⁴ <http://www6.inra.fr/cati-icat-atweb/Ontologies/Transmat> et <http://agroportal.lirmm.fr/ontologies/TRANSMAT>

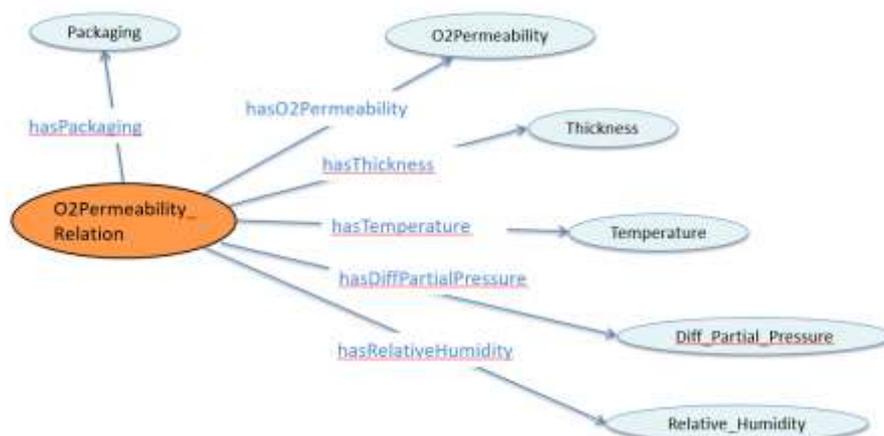


Figure 5. La relation n-aire O_2 Permeability Relation.

De manière plus générale, la RTO du logiciel @Web a été définie pour représenter des relations n-aires entre des données quantitatives expérimentales. Nous avons choisi de représenter des relations n-aires sans arguments différenciés telles que recommandé par le W3C (Noy et Rector, 2006), ce qui correspond au cas le plus général d'utilisation des relations n-aires. De plus nous avons choisi d'utiliser le patron qui représente une relation n-aire à l'aide d'un concept, relié à ses arguments par des propriétés. Nous présentons dans la suite la composante conceptuelle de la RTO, puis sa composante terminologique (voir **Figure 6**). La composante conceptuelle de la RTO est composée de deux parties: une ontologie noyau qui permet de représenter de manière générique (i.e. indépendamment d'un domaine particulier) des relations n-aires et une ontologie de domaine qui permet de représenter les concepts spécifiques à un domaine donné.

Nous avons ainsi découpé l'ontologie noyau en deux sous-parties : une partie supérieure, appelée ontologie noyau supérieure, qui permet de représenter des relations n-aires entre n'importe quels arguments, et une partie inférieure, appelée ontologie noyau inférieure, qui permet de représenter des relations n-aires entre des données quantitatives expérimentales. Dans l'ontologie noyau supérieure, les concepts génériques *Relation_Concept* et *Argument* permettent de représenter respectivement les relations n-aires et leurs arguments. Dans l'ontologie noyau inférieure, les concepts génériques *Dimension*, *UM_Concept*, *Unit_Concept* et *Quantity* permettent de gérer les quantités et leurs unités de mesure : le concept générique *Dimension* contient les dimensions qui permettent aux quantités et à leurs unités de mesures associées d'être classées par dimension et le concept générique *UM_Concept* contient les concepts qui permettent de gérer les conversions entre unités de mesure. Le concept générique *Symbolic_Concept* permet, quant à lui, de représenter les autres arguments non numériques des relations.

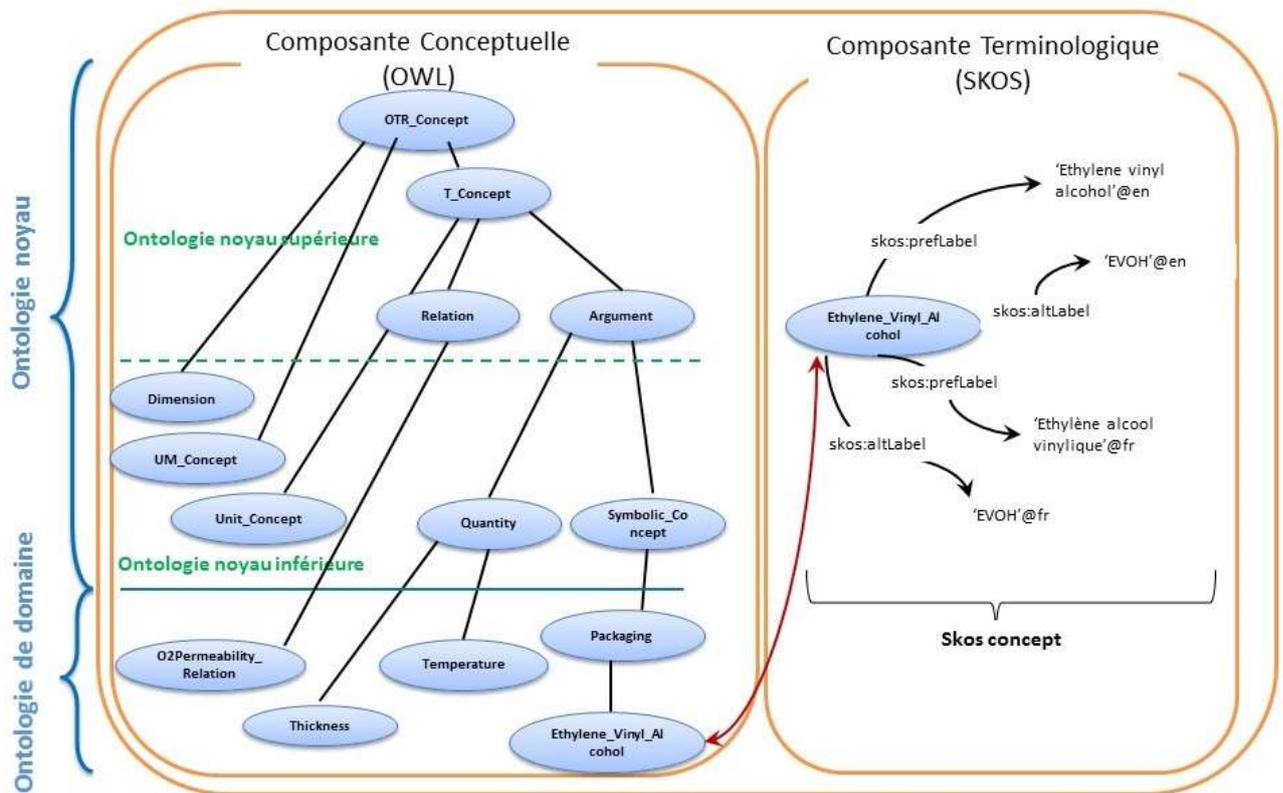


Figure 6. Les deux composantes d'une RTO.

L'ontologie de domaine contient les concepts spécifiques à un domaine d'application particulier. Ils apparaissent dans la RTO comme des sous-concepts des concepts génériques de l'ontologie noyau. La Figure 6 présente un extrait de la version actuelle de la hiérarchie des concepts de la RTO de domaine *Matter Transfer* qui contient 1145 concepts *Symbolic_Concept*, 77 concepts *Quantity*, 220 instances du concept *Unit_concept* et 57 relations n-aires (i.e. concepts *Relation_Concept*).

La composante terminologique de la RTO contient les termes multilingues (e.g. synonymes, abréviations) associés aux concepts de la composante conceptuelle.

Le logiciel @Web est un outil générique de collecte et d'annotation de données dans un entrepôt de données ouvert sur le Web qui s'appuie sur une ontologie noyau. Pour chaque nouveau domaine étudié, seules l'ontologie de domaine et sa composante terminologique devront être redéfinies pour bénéficier de l'ensemble des fonctionnalités de l'outil :

- exportation de modèles de fichiers CSV⁵ structuré selon un ensemble de relations n-aires sélectionnées dans l'ontologie de domaine ;
- importation de fichiers CSV contenant des données bien structurées et génération automatique de tableaux annotés stockés dans l'entrepôt de données ;
- association d'une URL et d'un DOI à chaque tableau annoté. L'URL permet l'accès au tableau par le Web. Le DOI est un mécanisme d'identification de ressource pérenne qui permet d'associer des métadonnées au tableau et à son URL qui peut changer au cours du temps ;
- évaluation de la fiabilité des données annotées à partir d'un ensemble de métadonnées ;

⁵ CSV: Comma-separated values

- interrogation multicritères de l'entrepôt de données contenant les tableaux annotés via une interface graphique dédiée ou un SPARQL⁶ endpoint ;
- accès en « open access » ou privé à l'entrepôt de données ;
- exportation des tableaux annotés au format CSV et RDF⁷.

Un exemple de réutilisation de données pour l'aide à la sélection d'emballages alimentaires

Une fois les données collectées dans l'entrepôt de données d'@Web, il est possible de les réutiliser. Nous présentons dans cette section un exemple de réutilisation des données dans un outil d'aide à la décision (OAD) qui permet de comparer des matériaux d'emballage. Cet OAD a été implémenté comme une application Web accessible sur la plateforme logicielle PLASTIC du département CEPIA (<http://pfl.grignon.inra.fr/EcoBioCapQuerying/>). Cet OAD permet de sélectionner un emballage alimentaire sous atmosphère modifiée en répondant à une question du type : « Je veux un emballage alimentaire qui préserve mon produit, fait à partir de ressources renouvelables, sans OGM, si possible transparent avec un coût inférieur à 3 € / kg ». Plusieurs courtes vidéos de démonstration de cet OAD sont disponibles au téléchargement⁸. La **Figure 7** présente les composants logiciels et bases de données qui permettent l'implémentation des fonctionnalités désirées de l'OAD. L'OAD présenté dans la **Figure 7** s'appuie sur deux bases de données : (i) une base de données « Produits frais » contenant les paramètres respiratoires des produits frais utilisés par le module de simulation de l'atmosphère modifiée de l'emballage et, (ii) une base de données « Emballages » décrivant notamment les caractéristiques suivantes des emballages : les perméabilités O₂/CO₂, l'épaisseur du film, les propriétés mécaniques, le prix, la biodégradabilité, la transparence. Ces bases de données sont alimentées à partir des données de la littérature et des fiches fournisseurs pour les emballages commercialisés. Les données des deux bases sont extraites de l'entrepôt de données d'@Web et stockées dans une base de données relationnelle MySQL (www.mysql.com). À partir des propriétés du produit frais à emballer, le modèle de simulation EAM permet de calculer les perméabilités optimales de l'emballage cible. Ces perméabilités optimales ainsi que d'autres préférences sur les caractéristiques de l'emballage (e.g. la transparence, la biodégradabilité) sont utilisées par un moteur d'interrogation multicritères bipolaire flexible afin de renvoyer les emballages de la base les plus pertinents (Guillard et al., 2015).

⁶ SPARQL: SPARQL Protocol and RDF Query Language

⁷ RDF: Resource Description Framework

⁸ umr-iate.cirad.fr/axes-de-recherche/ingenierie-des-connaissances/themes-de-recherche/ecobiocap-dss

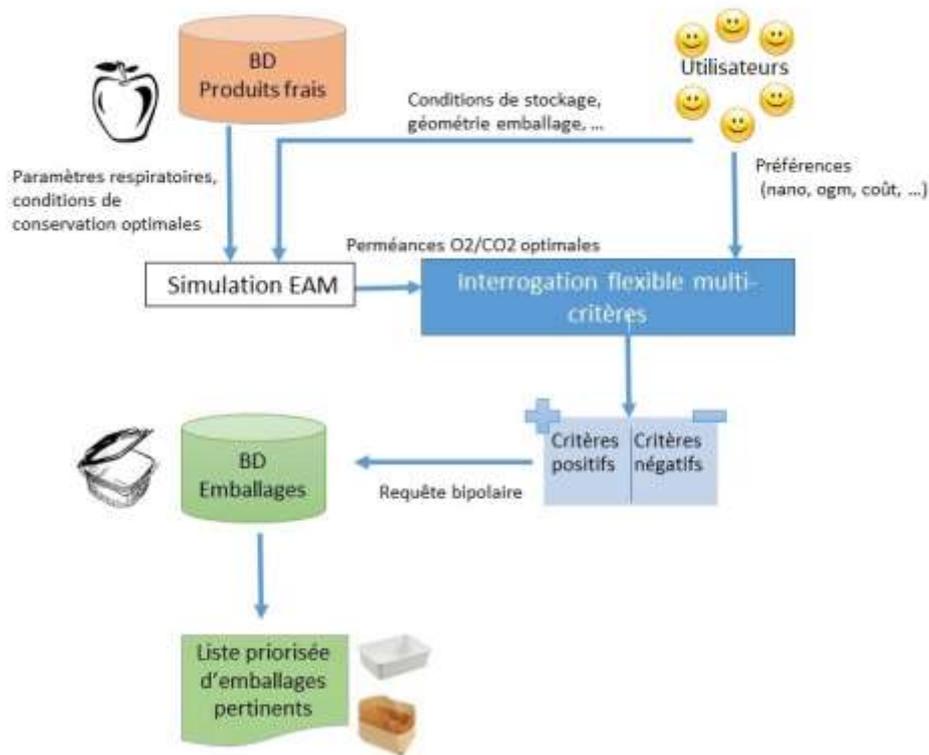


Figure 7. Architecture de l'OAD.

Dans ce qui suit, nous présentons un scénario d'utilisation de l'OAD pour comparer des matériaux d'emballages. La Figure 8 présente l'extrait d'un fichier CSV contenant des données de perméabilité structurées selon la RTO de domaine *Mass Transfer*. Une fois annotées et collectées dans l'entrepôt de données d'@Web, ces données sont répliquées dans la base « Emballages » de l'OAD.

	H	I	K	S	T	U	V	W	AA	
1	AUTHORS	YEAR	O2 Permeability_MIN	O2 Permeability_MAX	O2 Permeability_UNIT	Temp	Temp	Temp	Packaging	Packaging_ORIGINALVALUE
2	K. Syverud et P. Stenius	2009	17	18.5	cm3.m-2.day-1	23	23 °C		Cellulose	Microfibrillar cellulose (MFC) A
3	K. Syverud et P. Stenius	2009	17	17	cm3.m-2.day-1	23	23 °C		Cellulose	Microfibrillar cellulose (MFC) C
4	K. Syverud et P. Stenius	2009	50	130	cm3.m-2.day-1	23	23 °C		Polyester	Polyester oriented
5	K. Syverud et P. Stenius	2009	9	15	cm3.m-2.day-1	23	23 °C		Polyester	Polyester oriented, PVdC coated
6	K. Syverud et P. Stenius	2009	7800	7800	cm3.m-2.day-1	23	23 °C		Polyolefin	Polyethylene LD
7	K. Syverud et P. Stenius	2009	2600	2600	cm3.m-2.day-1	23	23 °C		Polyolefin	Polyethylene HD
8	K. Syverud et P. Stenius	2009	3	3	cm3.m-2.day-1	23	23 °C		Cellophane Film	Cellophane
9		2003	1078	1078	amol.m-1.s-1.Pa-1	20	20 °C		Polyolefin	Low-density polyethylene film
10	M. D. Sanchez-Garcia, E. Giménez	2008	1.5E-19	2.1E-19	m3.m-2.s-1.Pa-1	24	24 °C		Polyester	NanoterPHB -- 4% of wt (Nanoter TM 2212) based on K
11	M. D. Sanchez-Garcia, E. Giménez	2008	3.9E-19	4.1E-19	m3.m-2.s-1.Pa-1	24	24 °C		Polyester	NanoterPHB-Blend -- 1% of wt (Nanoter TM 2212) base
12	M. D. Sanchez-Garcia, E. Giménez	2008	3.8E-19	4E-19	m3.m-2.s-1.Pa-1	24	24 °C		Polyester	NanoterPHB-Blend -- 1% of wt (Nanoter TM 2212) base
13	M. D. Sanchez-Garcia, E. Giménez	2008	2.1E-19	2.7E-19	m3.m-2.s-1.Pa-1	24	24 °C		Polyester	NanoterPHB-Blend -- 4% of wt (Nanoter TM 2212) base
14	M. D. Sanchez-Garcia, E. Giménez	2008	2.6E-19	3E-19	m3.m-2.s-1.Pa-1	24	24 °C		Polyester	NanoterPHB-Blend -- 4% of wt (Nanoter TM 2212) base
15	M. D. Sanchez-Garcia, E. Giménez	2008	5.8E-18	5.8E-18	m3.m-2.s-1.Pa-1	24	24 °C		Polyester	PCL -- Polycaprolactone grade FB100
16	M. D. Sanchez-Garcia, E. Giménez	2008	2.298E-19	2.302E-19	m3.m-2.s-1.Pa-1	24	24 °C		Polyester	PHB -- Material with density 1.25 g/cm3 is a met-proce
17	M. D. Sanchez-Garcia, E. Giménez	2008	4.1995E-19	4.2005E-19	m3.m-2.s-1.Pa-1	24	24 °C		Polyester	PHB blend -- PHB/PCL, 80:20 wt/wt
18	M. D. Sanchez-Garcia, E. Giménez	2008	5.196E-19	5.204E-19	m3.m-2.s-1.Pa-1	24	24 °C		Polyester	PHB blend -- PHB/PCL, 80:20 wt/wt
19	M. D. Sanchez-Garcia, E. Giménez	2008	3.3E-19	3.3E-19	m3.m-2.s-1.Pa-1	24	24 °C		Polyethylene Terephthalate	PET -- PET commercial
20	M. D. Sanchez-Garcia, E. Giménez	2008	3.8E-19	3.8E-19	m3.m-2.s-1.Pa-1	24	24 °C		Polyethylene Terephthalate	PET -- PET commercial

7800	7800	cm3.m-2.day-1	23	23 °C	Polyolefin	Polyethylene LD
------	------	---------------	----	-------	------------	-----------------

Figure 8. Exemple de données annotées dans @Web.

L'interface graphique de l'OAD est composée de trois parties :

- la partie supérieure, présentée dans la **Figure 9**, permet de calculer la perméance optimale pour l'emballage ciblé à partir des paramètres caractérisant l'aliment, les conditions de stockage et la géométrie de l'emballage. L'interface se connecte à la base de données « Produits frais » pour récupérer les caractéristiques associées à l'aliment sélectionné. La **Figure 9** présente les perméances optimales O₂/CO₂ calculées par l'OAD pour une durée de vie de 7 jours, à une température ambiante de 20°C, une masse de produit de 500 g, un volume de l'emballage de 1 litre et une surface de 756 cm² ;
- la partie centrale de l'interface permet à l'utilisateur d'exprimer ses préférences. Dans cette version de l'OAD, il peut uniquement exprimer ses préférences sur les perméances O₂/CO₂, la température de stockage, la biodégradabilité et la transparence de l'emballage. La requête présentée dans la **Figure 9** peut s'exprimer de la manière suivante : « Je veux un emballage adapté à la conservation d'un abricot Bergeron pendant 7 jours (valeurs optimales de perméance O₂/CO₂ calculées par l'OAD) pour un intervalle de température variant entre 14 et 26°C ». Il est à noter que les valeurs optimales de perméance O₂/CO₂ calculées par l'OAD, sont automatiquement recopiées dans la partie centrale de l'interface avec un écart type prédéfini pour l'intervalle min-max et l'intervalle min-max élargi. Ces valeurs peuvent être modifiées par l'utilisateur avant de soumettre la requête au moteur d'interrogation flexible qui compare ces valeurs aux caractéristiques des emballages stockés dans la base « Emballages » ;
- la partie inférieure est dédiée à l'affichage du résultat de la requête, présentée dans la **Figure 10**, dans le cas de l'abricot. L'emballage classé en première position est un Polyéthylène HD.

EcoBioCap - Optimize permeabilities

Food properties

Apricot Bergeron

Mass (kg): 0.5

Shelf life (day): 7

Temperature (°C):

Optimal atmosphere value:

O₂ (%): 3

CO₂ (%): 2

Respiration properties:

RRO₂ max (mmole/kg/h): 0.415

RQ (RRCO₂ / RRO₂): 0.78

KmO₂ (Pa): 4500

KICO₂ (Pa): -1

Packaging geometry

Surface (cm²): 756

Volume (l): 1

run simulation

clear

Permeance O₂ (mol.m-2.s-1.Pa-1) 1.411684e-11

Permeance CO₂ (mol.m-2.s-1.Pa-1) 1.29492e-10

Permeability O₂ (mol.m-1.s-1.Pa-1 - 50 µm) 7.058419e-16

Permeability CO₂ (mol.m-1.s-1.Pa-1 - 50 µm) 6.474602e-15

Preferences associated with criteria

allow the ranking of packagings with unknown values for mandatory criteria

	enlarge min	min	max	enlarge max	mandatory	optional
O ₂ permeance	9.881786e-12	1.270515e-11	1.552852e-11	1.835189e-11	<input checked="" type="checkbox"/>	<input type="checkbox"/>
CO ₂ permeance	9.064443e-11	1.165428e-10	1.424412e-10	1.683397e-10	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Temperature	14	18	22	26	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Biodegradability					<input type="checkbox"/>	<input type="checkbox"/>
Transparency	<input type="text" value="transparent"/> <input type="text" value="translucent"/> <input type="text" value="opaque"/>				<input type="checkbox"/>	<input type="checkbox"/>

rank packagings

Figure 9. Interface graphique de l'OAD présentant les valeurs de perméance optimales obtenues pour le cas d'un abricot.

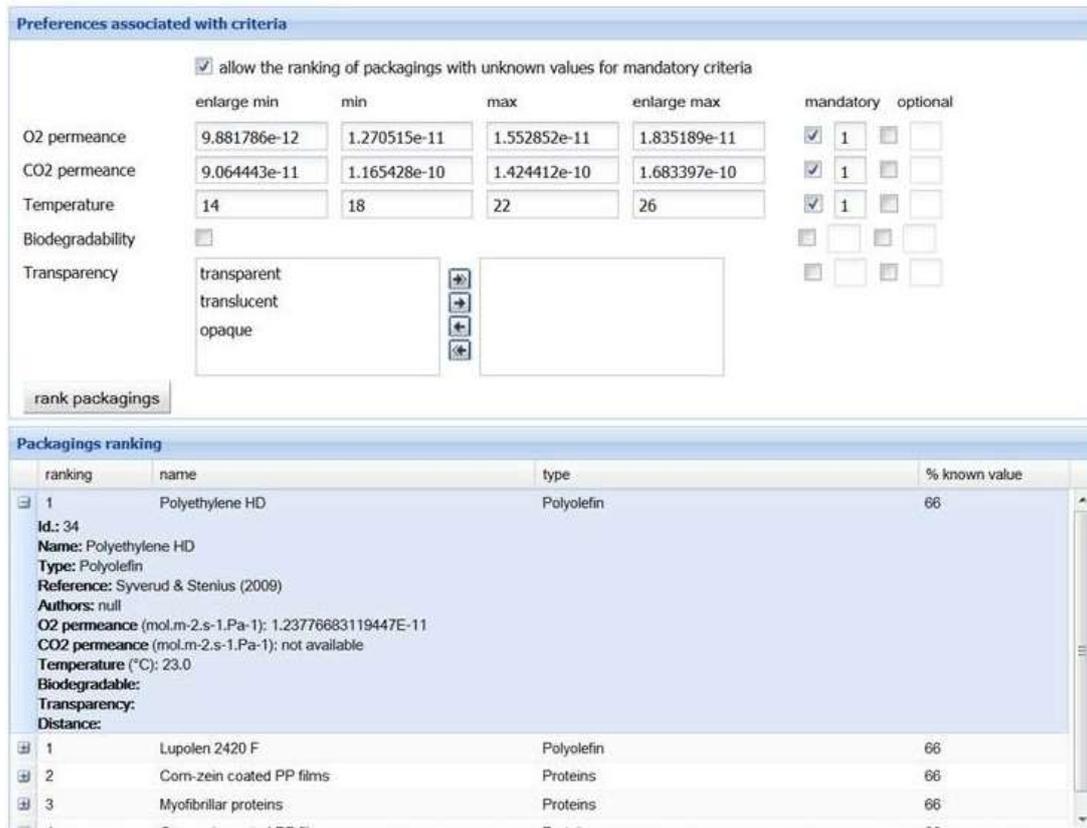


Figure 10. Interface graphique de l'OAD présentant une liste triée d'emballages par ordre de pertinence décroissante. Des informations détaillées sont affichées pour le premier emballage.

Augmenter la visibilité du travail du chercheur par la valorisation des données produites

L'ouverture des données dans une démarche de type Open Science permet d'accroître la visibilité du travail du chercheur. Celui-ci a la possibilité de publier ses jeux de données dans un entrepôt ouvert et de publier un data paper⁹ en plus des publications classiques dans des revues ou journaux. Chacun des jeux de données d'un chercheur annoté dans l'entrepôt de données d'@Web étant pourvu d'une URL et d'un DOI, ces trois types de productions peuvent se citer respectivement comme indiqué dans la Figure 11.

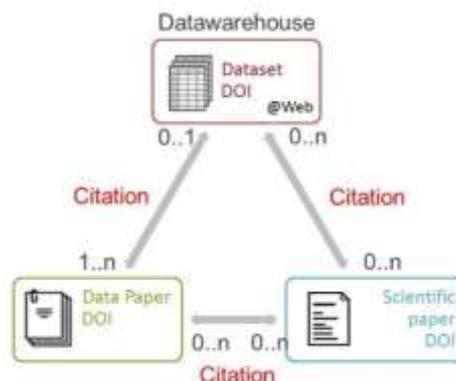


Figure 11. Citations entre les trois types de production d'un chercheur.

⁹ <https://www6.inra.fr/datapartage/Partager-Publier/Valoriser-ses-donnees/Publier-un-Data-Paper>

Nous illustrons ce processus de communication élargie au travers d'un exemple de data paper (Guillard et al., 2016) qui référence des tableaux annotés dans l'entrepôt de données d'@Web (cf **Figure 12**). Les données sont accessibles dans l'entrepôt d'@Web à partir de l'URL associée à chaque tableau annoté.

I. Data

Data shared with this article are more than 100 data of solubility and diffusivity of gases (O₂ and CO₂) in food samples. These data are stored in a data warehouse called @Web in which the data management is guided by ontology.

All data are available for uploading at the URL specified below and recalled in the table hereafter with the details about the nature and amount of data available at each URL.

Data type	Table URL (copy/paste the URL in your Internet browser)	Amount of data
CO ₂ solubility	pfl.grignon.inra.fr/atWeb/TableServlet?viewTable=2775&idDoc=1335&id=35272672	34
	pfl.grignon.inra.fr/atWeb/TableServlet?viewTable=2776&idDoc=1335&id=35305550	21
	pfl.grignon.inra.fr/atWeb/TableServlet?viewTable=2773&idDoc=1335&id=35245144	48
	pfl.grignon.inra.fr/atWeb/TableServlet?viewTable=2732&idDoc=1332&id=34354344	3
	pfl.grignon.inra.fr/atWeb/TableServlet?viewTable=2780&idDoc=1346&id=35361064	12
CO ₂ diffusivity	pfl.grignon.inra.fr/atWeb/TableServlet?viewTable=2826&idDoc=1346&id=36350532	11
	pfl.grignon.inra.fr/atWeb/TableServlet?viewTable=2779&idDoc=1346&id=35346176	11
	pfl.grignon.inra.fr/atWeb/TableServlet?viewTable=2778&idDoc=1346&id=35333312	12
	pfl.grignon.inra.fr/atWeb/TableServlet?viewTable=2777&idDoc=1346&id=35320430	16

Figure 12. Extrait de data paper.

Il est possible d'associer un DOI à un tableau annoté dans l'entrepôt d'@Web comme illustré dans la **Figure 13**. Il faut pour cela renseigner plusieurs métadonnées du tableau (e.g. titre, auteurs, mots-clés, URL du tableau). Dans l'exemple de la **Figure 13**, la métadonnée *packaging O2 permeability* a été associée au *sujet* du tableau. Ce tableau pourra être retrouvé par une interrogation sur le mot-clé *packaging O2 permeability* en utilisant le moteur de recherche DataCite comme illustré dans la **Figure 14** et la **Figure 15**.

Manage DOI ✖

DOI Metadata for document's table

Title * :

Publisher * :

Authors * :

Subject / Keywords :

Landing Page* :

Ressource type * :

Figure 13. Association d'un DOI à un tableau annoté d'@Web.

Figure 14. Exemple de requête Datacite.

Figure 15. Exemple de réponse Datacite.

Conclusion, freins, leviers

Nous avons présenté dans cet article une démarche Open Science qui permet de collecter dans un entrepôt de données ouvert sur le Web des données provenant de sources hétérogènes. Cette démarche met en œuvre les principes FAIR de gestion des données (i.e. Findable, Accessible, Interoperable, Reusable), le principe « Findable » est réalisé en proposant au travers d'@Web la création de DOI rendant les métadonnées des tableaux de données visibles et indexables par les moteurs de recherche du Portail data INRA et de DataCite. Le principe « Accessible » est satisfait par le stockage de manière pérenne sur le Portail data INRA des métadonnées, référençant elles-mêmes les données disponibles sur @Web dans des formats ouverts¹⁰ (principe « Interoperable ») et réutilisables sous licence CC-BY pour d'autres applications comme la réalisation d'un outil d'aide à la sélection d'emballages alimentaires (principe « Reusable »).

¹⁰ CSV et RDF

Les tableaux de données d'@Web seront bientôt répliqués sur le Portail Data INRA. La mise à disposition des données de la recherche et la valorisation des données sous forme de data paper peut être vue comme une injonction paradoxale pour le chercheur à qui l'on demande de valoriser son travail sous forme de publications scientifiques originales pour lesquelles les données produites sont la matière première. La mise en place d'outils facilitant le travail de mise à disposition des données de la recherche (e.g. Portail Data INRA, Open Air+) et la reconnaissance institutionnelle de la production de data papers devraient favoriser ce nouveau mode de diffusion des résultats de la recherche.

Références bibliographiques

Buche P, Dibie-Barthélemy J, Ibanescu L, Soler L (2013) Fuzzy Web Data Tables Integration Guided by an Ontological and Terminological Resource. *IEEE Trans. Knowl. Data Eng.* **25**(4): 805-819.

Guillard V, Buche P, Destercke S, Menut L, Guillaume C, Gontard N (2015) A Decision Support System for designing biodegradable packaging for fresh produce. *Comptu Electron Agric.* **111**:131-139.

Guillard V, Buche P, Dibie J, Dervaux S, Acerbi F, Chaix E, Gontard N, Guillaume C (2016) CO₂ and O₂ solubility and diffusivity data in food products stored in data warehouse structured by ontology. *Data in brief* **7**:1556-1559

Noy N, Rector A. (2006) Defining N-ary Relations on the Semantic Web.
Consulté sur : <http://www.w3.org/TR/swbp-n-aryRelations/>

Cet article est publié sous la licence Creative Commons (CC BY-SA).



<https://creativecommons.org/licenses/by-sa/4.0/>

Pour la citation et la reproduction de cet article, mentionner obligatoirement le titre de l'article, le nom de tous les auteurs, la mention de sa publication dans la revue « Le Cahier des Techniques de l'INRA », la date de sa publication et son URL).