Relating transformation process, eco-design, composition and sensory quality in cheeses using PO^2 ontology

Caroline Pénicaud [a, *], Liliana Ibanescu [b], Thomas Allard [c], Fernanda Fonseca [a], Stéphane Dervaux [b], Bruno Perret [a, d], Hervé Guillemin [d, e], Solange Buchin [e], Christian Salles [c], Juliette Dibie [b], Elisabeth Guichard [c]

[a] UMR GMPA, AgroParisTech, INRA, Université Paris-Saclay, 78850, Thiverval-Grignon, France
[b] UMR MIA-Paris, AgroParisTech, INRA, Université Paris-Saclay, 75005, Paris, France
[c] Centre des Sciences du Goût et de l'Alimentation, AgroSup Dijon, CNRS, INRA, Université Bourgogne Franche-Comté, F-21000 Dijon, France
[d] PLASTIC Platform, INRA, 78850, Thiverval-Grignon, France
[e] URTAL, INRA, 39800 Poligny, France

ARTICLE INFO

ABSTRACT

To provide answers to sustainability challenges, a database called BaGaTel, guided by the PO^2 ontology, has been built to integrate data to reformulate dairy products taking into account nutritional and sensory properties together with environmental concerns. In this paper, BaGaTel was queried to address questions dealing with the eco-design of hard cheese processing, in relation to composition, sensory quality and rheological properties. For the formulation of hard cheese, BaGaTel made it possible to estimate missing data in a dataset supposing that samples have common characteristics. For environmental concerns, BaGaTel gave hints about relevant data that need to be acquired and made possible the estimation of missing data. The common vocabulary and structure provided by the PO^2 ontology allowed combining and integrating into BaGaTel data from different projects, giving relevant answers to different questions, and therefore proving its suitability as a support tool for multi-criteria assessment of food systems.

© 2019 Elsevier Ltd. All rights reserved.

1. Introduction

The food processing sector is facing sustainability challenges of growing complexity, such as global warming, the increase in overweight, obesity or population ageing. These problems, while having a tremendous impact on populations, urge the food industry to develop new strategies to formulate well-balanced products in terms of nutritional requirements (e.g., less fat, sugar and salt), animal/vegetable protein supply and sensory acceptability by consumers, while using eco-friendly transformation processes. In this paper, we will focus on dairy products, which have been reported to have a high environmental impact (Weidema et al., 2008) and more specifically cheeses. Cheeses have a variety of compositions (e.g., different fat and protein types and content), structures (i.e., from liquid to solid) and micro-structures (e.g., fat droplets,

protein networks), which open up many reformulation opportunities for their production. However, even if a lot of studies have been conducted on dairy products, considering on one hand the environmental impact of food processing and on the other the nutritional and sensory quality of foods, there is no integrated study combining all these data.

To tackle this challenge, data and knowledge from different domains in science (e.g., nutrition, sensory perception, eco-design, microbiology, biochemistry, process engineering) need to be combined with data and knowledge from environmental analysis.

The environmental impact of food and food processing is seriously questioned. Food products have been reported to contribute to 20−30% of European environmental impact (Tukker et al., 2006). Life Cycle Assessment (LCA) is a standard method (ISO 14040:2006; ISO, 2006) that allows quantification of the environmental impacts of a product, process or service during its whole life cycle, and it is widely applied to food systems (Roy et al., 2009) and more specifically on semi-hard cheeses (Berlin, 2002). However, environmental impacts are more frequently related to the mass of food

* Corresponding author. Tel.: +331 30 81 54 17.
  E-mail address: caroline.penicaud@inra.fr (C. Pénicaud).

produced than to its quality, which can lead to ineffective and even counterproductive eco-design process options (Igos et al., 2013; Pénicaud, Monclus, Perret, Passot, & Fonseca, 2018). As an example, Berlin (2002) suggested that the environmental impact of semi-hard cheese should be reduced by decreasing the wastage of the milk, but this could affect quality of the cheese.

There is thus a crucial need to take into account relevant parameters of food quality in any environmental impact assessment. Product quality can be assessed by taking into account its sensory quality, for instance. Sensory quality is the result of all the sensory modalities (i.e., aroma, taste and texture) perceived during the physiological process necessary for in-mouth food breakdown and varies with product formulation and technology. Relationships among food composition, structure, sensory properties and food breakdown in the mouth have already been investigated by in vivo and/or in vitro studies conducted on dairy products (Doyennette et al., 2011; Guichard, Repoux, Qannari, Laboure, & Feron, 2017; Lauverjat, Deleris, Trelea, Salles, & Souchon, 2009; Lawrence et al., 2012). However, these studies were conducted on different dairy food products, did not measure the same parameters and did not take into account environmental parameters. It is thus difficult from these data to link quality assessment to environmental impact. The obtained data are also difficult to compare because of (i) their heterogeneity (e.g., qualitative data such as flavour perception versus quantitative or semi-quantitative data such as energy consumption) (ii) their measurement scales (e.g., microstructure of product vs impact of the whole system on global warming), and (iii) their evolution over time: some measured properties may change over time through the different unit operations for production (e.g., content in unsaturated lipids), whilst others do not (e.g., process working conditions).

In addition, data are generally acquired during a specific research project, for a given food product and are focused on specific parameters. Because of natural variability of food products, it is hard to directly compare two products coming from different projects. Moreover, one project may not produce a great number of measurements: a standard experimental procedure is usually based on triplicates, which is relatively low from a statistical point of view. Besides, data coming from a specific project are generally stored in laboratory books and/or on personal computers of the persons involved in the project, and not widely shared with the scientific community.

The great challenge is thus to combine, aggregate and integrate all data and knowledge that have been collected in different projects covering specific domains. Data heterogeneity, their multiscales dimension, sparseness and their evolution over time through different unit operations of the production process represent important bottlenecks. Moreover, expert knowledge is often implicit and difficult to acquire. Nevertheless, data and knowledge storage and integration enriched with the bibliographic data is of utmost importance for sustainable process development.

To address data and knowledge integration, a relevant solution is the use of an ontology (Doan, Halevy, & Ives, 2012). Ontology is defined as a formal common vocabulary of a given domain, shared by the domain experts. Ontology is designed to represent the knowledge from one domain in terms of concepts, relationships among these concepts and instances of these concepts (Guarino, Oberle, & Staab, 2009). Moreover, publishing ontologies on the Linked Open Data (LOD) cloud (http://linkeddata.org) and building networks of interconnected ontologies (Suárez-Figueroa, Gómez-Pérez, & Fernández-López, 2012) facilitate data integration and data sharing, such as giving access to data from specific disciplines or data produced within specific geographic regions (Bizer, 2013).

In the food and human nutrition domain, eleven ontologies are listed on the portal providing the map of agri-food data standards (http://vest.agrisemantics.org), a recent initiative within a GODAN (Global Open Data for Agriculture and Nutrition) Action project (http://www.godan.info). For instance, Food Track & Trace Ontology has been developed for traceability purposes, with the aim of being connected with a global track and trace information system (Pizzuti, Mirabelli, Sanz-Bobi, & Gomez-Gonzalez, 2014). This ontology has been extended into MESCO (Meat Supply Chain Ontology) for supporting the management of meat traceability along the whole supply chain, from farmer to final consumer (Pizzuti, Mirabelli, Grasso, & Paldino, 2017). Another example is the OFPE (Ontology for Food Processing Experiment; Muljarto, Salmon, Neveu, Charnomordic, & Buche, 2014) dedicated to food processing experiments and reused in AFEO (Agri-Food Experiment Ontology; Muljarto et al., 2017) to guide data integration in viticulture and winemaking. Moreover, close to environmental concern, a decision support system has been created to allow the use of data found in the literature to assess environmental sustainability of biorefinery systems (Lousteau-Cazalet et al., 2016). This couples ontologies, soft computing techniques and environmental factor computation. However, none of these ontologies was suited to represent a food transformation process described by a set of experimental observations available at different scales and changing over time through the different unit operations of the production process. Only such an ontology would permit handling of multi-criteria indicators associating the quality of the food product, the transformation process constraints and the environmental impact of the whole system.

With this objective, our collaborative network gathering scientists in food process, oral physiology and sensory perception, ecodesign and computer science, built a first ontology for the ecodesign of transformation processes (Dibie, Dervaux, Doriot, Ibanescu, & Pénicaud, 2016). This ontology has been extended to PO^2, a process and observation ontology in food science (available at http://agroportal.lirmm.fr/ontologies/PO2/), to integrate data for the formulation of dairy products, taking into account nutritional and sensory properties (Ibanescu, Dibie, Dervaux, Guichard, & Raad, 2016). A database, called BaGaTel, structured by the PO^2 ontology has been built, hosted by the PLASTIC platform of INRA (http://www.pfl-cepia.inra.fr). Data from different sources (e.g., experiments, papers, databases) are manually annotated with concepts and relations from PO^2 ontology and are thus stored into specific EXCEL files that were previously designed according to the structure of PO^2. Data from EXCEL files are uploaded into the BaGaTel database. Once the ontology has been built, the specific vocabulary has been fixed, and the data and knowledge collected have been integrated in a common database, the question of their querying and analysis for decision support issues can be addressed.

The database and the vocabulary of the ontology have been designed for two complementary applications. The first one is to allow researchers in food science to answer research questions. The second is to help food producers formulate food products that answer different constraints (nutritional, sensory, and environmental). The structure of the database and the vocabulary had to be suitable to facilitate the querying of data by both researchers and food producers. The objective of this paper is to demonstrate the ability and the suitability of the developed tools to provide relevant data in response to questions which involve multi-criteria indicators. Several questions in the context of formulation and environmental concerns are addressed.

In Section 2, this paper presents the different data available on dairy products in BaGaTel database, and more specifically on hard cheeses, with a focus on the parameters used to formulate the requests. Section 2 also describes briefly PO^2 ontology. Section 3 reports on how the use of BaGaTel database structured according to PO^2 allows answering to specific questions relating process, product quality and eco-design parameters.

2. Materials and methods

2.1. Hard cheeses

Among dairy products, cheese is widely consumed, especially in Europe and North America, and is the main growth product within the dairy sector (Fox, 2011). Cheese is a fermented food derived from milk. Although cheese making is a linear process, it involves many manufacturing steps and at each step, different parameters can be measured for production control. This makes cheese a very interesting model food product to study to understand and qualify the links among manufacturing parameters, product quality and environmental impact. Considering the fact that a great number of different types of cheese exists, due to the diversity of composition, processes and ripening steps, the present paper will focus on hard cheeses and related model systems.

2.1.1. Available data and addressed questions

There are a significant number of research projects conducted on hard cheeses, from which data were successfully imported into the BaGaTel database following the PO^2 structure. Available data on hard cheeses including technical process descriptions and parameters, composition and structure of the studied cheese during different process steps and sensory perception parameters, were collected during the following projects: TRUEFOOD (EU-FP6-FOOD-16264: 2006−2010; https://cordis.europa.eu/result/rcn/46734_en.html); SmartRipe (EU-FP7-KBBE-613827: 2013−2015; https://cordis.europa.eu/project/rcn/111141_fr.html); and PRASEL (Programme de recherche en alimentation sur le sel. French national project 2005−2008). Summaries of TRUEFOOD, SmartRipe and PRASEL projects are provided as Supplementary material files SF1, SF2 and SF3, respectively. Complementary data were imported into the field of in-mouth food breakdown, aroma and taste compounds release and sensory acceptability from projects done on model cheeses similar to hard cheeses (Bigaski Ribeiro et al., 2016; Boisard et al., 2013, 2014; Feron et al., 2014; Gierczynski, Labouré, Sémon, & Guichard, 2007; Guichard et al., 2017; Lawrence et al., 2011; Mosca, Andriot, Guichard, & Salles, 2015; Phan et al., 2008; Tarrega, Yven, Sémon, & Salles, 2008, 2011).

These projects covered a wide range of cheese composition (e.g., lipid, protein, salt, microorganisms), technical processes from very simple model systems to real cheese production, cheese structure at different scales, sensory perception (i.e., aroma, taste and texture), in-mouth breakdown (e.g., masticatory parameters, food bolus structure), as well as in vitro and in vivo aroma and taste compounds release.

Other data from a process eco-design project (CellExtraDry; Eco-design preservation processes of functional yeast enriched with antioxidants. Funded by ANR-IC-Qualiment) were also imported into BaGaTel, covering inputs and outputs used for the transformation processes (e.g., consumptions and wastes of energy, water, steam) of stabilised micro-organisms, together with quality data for yeasts (e.g., enumeration, antioxidant capacity), the processing conditions (e.g., temperature, pressure) and resulting environmental impacts computed by LCA. The summary of CellExtraDry project is provided as Supplementary material file SF4. Energy consumption for processes and processing conditions in the case of hard cheeses are also available from TRUEFOOD (EU-FP6-FOOD-16264: 2006−2010) and SmartRipe (EU-FP7-KBBE-613827: 2013−2015) projects.

As not all parameters were measured at the same time on the same cheese sample, the aim of this study is to investigate how and to what extent it is possible to benefit from data integrated from different projects to answer to specific questions. Questions focus on relationships among composition, sensory quality and rheological properties of hard cheeses, and also on their environmental impact.

The first three questions addressed the problem of formulation by combining data about cheese composition during production, sensory quality and rheological parameters. In fact, sensory description and rheological parameters are commonly used in the literature to assess the quality of cheese samples differing in composition and process (Foegeding, Brown, Drake, & Daubert, 2003; Hailu et al., 2018). The last two questions address the environmental impact issue of the production process. The questions were:

Q1. Which hard cheese sample has the highest value for "taste intensity"?
Q2. What is the composition (participant part) of this sample at different steps in the process?
Q3. What are the characteristics (observation part) measured on this sample?
Q4. Which parameters need to be measured throughout the process to perform a LCA of the sample?
Q5. What are the values of the relevant parameters (observation part) needed to conduct the LCA?

2.1.2. Hard cheese production process

Data imported into BaGaTel describe the production of French hard cheeses made from raw cows' milk in the Franche-Comté region of eastern France. The general outline of French hard cheese production steps can be described as follows (data from TRUEFOOD, 2008). The composition of the milk is analysed (i.e., lipid, protein, lactose and water content) during the *Milk reception* step. Then, a cream separator is used to separate the cream from the skim milk by turning the milk at high speed at the *Skimming* step. During the *Milk standardisation* step, cream and skimmed milk are combined to ensure a standard fat/protein ratio. The *Cooling* step consists of cooling the standardised milk in refrigerated tanks. After pumping into the manufacturing vat, the milk is inoculated with lactic starters during the *Inoculation of the milk* step. During the *Renneting* step, rennet is added to the milk causing milk gelation. Different treatments are then applied to the milk in the vat at the *Step in the vat*. To remove part of the whey trapped in the coagulated matrix, the curd is cut, then stirred and heated. Withdrawal of whey is favoured by increasing outlet surfaces and by retracting the matrix. In the *Draining* step, the mass of curd grains is separated from the whey, and put into moulds during the *Moulding* step. During the *Pressing* step, the remaining intergranular whey is evacuated and the cheese is given its final form. The *Demoulding* step consists in removing the cheese from the mould in which it has been set to take its specific shape. During the *Pickling* step, the cheese is immersed in an aqueous solution saturated with sodium chloride. Finally, the cheese is placed in three successive ripening cellars with different controlled temperatures and relative humidities where it is regularly rubbed with salt solution. During this *Cheese ripening* step, fermentations continue and transform the cheese.

2.1.3. Sensory description of cheese samples

Sensory taste and texture profiles of cheeses were carried out by different trained panels using conventional sensory profiling (AFNOR, 2009). The main attributes used in this paper to describe the texture are springiness, firmness, granularity, hardness and moisture, while the attributes used to describe the taste are taste intensity, salty, sour and sweet. Their definitions are given in Table 1.

2.1.4. Rheological measurements

Rheological properties of the cheeses were determined by a uniaxial compression test at constant displacement rate. Rheological measurements were made with a Texture Analyser TA XT2 (Stable Microsystems Ltd). During the test, the sample was

Table 1
The main attributes used during sensory profiling.

Sensory characteristic	Attribute	Definition or reference used
Texture	Springiness	Mechanical attribute related to the rapidity of shape recovery after the application of a deforming force (adapted from AFNOR, 2009)
Texture	Firmness	Mechanical textural attribute related to the force required to achieve a given deformation of the sample by teeth penetration (adapted from AFNOR, 2009)
Texture	Granularity	Geometrical textural attribute relating to the perception of the size, shape and amount of particles in a product (AFNOR, 2009)
Texture	Hardness	Mechanical textural attribute relating to the force required to achieve a given penetration, or breakage of a product (AFNOR, 2009)
Texture	Moisture	Surface textural attribute that describes the perception of water absorbed by or released from a product (AFNOR, 2009)
Taste	Overall taste intensity	Intensity of sensations perceived by the taste organ when stimulated by certain soluble substances (adapted from AFNOR, 2009)
Taste	Salty	Basic taste produced by dilute aqueous solutions of various substances such as sodium chloride (AFNOR, 2009)
Taste	Sour	Basic taste produced by dilute aqueous solutions of most acid substances (e.g., citric acid and tartaric acid) (AFNOR, 2009)
Taste	Sweet	Basic taste produced by dilute aqueous solutions of natural or artificial substances such as sucrose or aspartame (AFNOR, 2009)

compressed at a constant crosshead speed according to its main axis between two parallel plates. Samples were compressed to 80% maximum deformation at 0.8 mm s^{-1} between parallel plates. Using recorded force and displacement data, engineering stress and Cauchy strain were calculated. From these data, the Young modulus MD (kPa), fracture stress σf (kPa), fracture strain εf ($-$), and work to fracture Wf (kJ m^{-3}) were determined as described by Noël, Zannoni & Hunter (1996).

The Young modulus is an estimation of the apparent elastic modulus and characterises the elastic properties of the cheese. The work to fracture describes the total deformation work up to the fracture. Fracture stress and the work to fracture describe the mechanical resistance of the cheese. Fracture strain characterises the deformability of the cheese.

2.1.5. Life Cycle Assessment methodology

LCA can be divided in four steps: goal and scope definition, data inventory, impact assessment and interpretation (ISO, 2006).

Step 1: *Goal and scope definition.* This phase consists in defining the objectives of the study, system boundaries, functional unit, data needed, assumptions and limits. The goal of the LCA presented in this paper was to show to what extent the BaGaTel database can help LCA practitioners perform the data inventory step. A simplified LCA was thus performed on a specific sample to illustrate the contribution of BaGaTel in a LCA approach. The system boundaries exactly fitted the system described in section 2.1.2. The functional unit was the production of one cheese of 10.5 kg ripened for 120 days, corresponding to the use of 115 L of raw milk. The data needed mainly came from data included in databases, no complementary measurement was performed. Because of the production process of the sample under study, no allocation had to be assumed. Obviously, such assumptions on data and allocation can limit the reliability of the obtained results. Nevertheless, this paper aims to show how the BaGaTel database can be helpful in LCA studies, but does not claim to provide thorough LCA results. Such a simplified approach was sufficient to achieve our goal.

Step 2: *Data inventory.* This stage concerns quantification of the mass and energy flows used and emitted by the system under study. Modelling of the system requires data collection for the entire life cycle and this is often the most time-consuming stage of an LCA. Primary data are measured data, secondary data are obtained from calculations or databases (e.g., Ecoinvent).

The data inventory is detailed in the results section to show how the BaGaTel database can be useful to perform this step. Briefly, data used to perform the simplified LCA came from BaGaTel database, supplemented by Ecoinvent database.

Step 3: *Impact assessment.* Environmental impacts are evaluated by using identified inputs and outputs through specific indicators usually fixed in internationally recognised methods (e.g., ILCD). In this paper, environmental impacts were calculated by using ILCD

2011 Midpoint + method (European Commission, 2011). This method provides environmental impacts with regards to sixteen indicators (climate change, ozone depletion, human toxicity - non-cancer effects, human toxicity - cancer effects, particulate matter, ionising radiation HH, ionising radiation E (interim), photochemical ozone formation, acidification, terrestrial eutrophication, freshwater eutrophication, marine eutrophication, freshwater ecotoxicity, land use, water resource depletion, mineral, fossil and renewable resource depletion). However, for clarity reasons, only results on climate change, ozone depletion and acidification indicators are presented in the results section. System modelling and impact assessment were performed by using Simapro software (V8.4.0.0 PRé consultant).

Step 4: *Interpretation.* An LCA is an iterative process and the aim is to analyse results and hypotheses and discuss them at each Step 1, 2 and 3. This will help decision makers to make an improved choice. This paper aims at showing how the BaGaTel database can be helpful in LCA studies, but does not claim to provide thorough LCA results. As a consequence, LCA assumptions and results are not deeply discussed in this paper.

2.2. PO^2 ontology description

PO^2 ontology (Ibanescu et al., 2016) allows representation of a food transformation process described by a set of experimental observations available at different scales and changing over time through the different unit operations of the production process. Ontology is designed in two layers: a core layer and a domain layer. The core layer is composed of three main parts, presented in Fig. 1 (adapted from Ibanescu et al., 2016).The domain layer allows one to represent a specific domain, here the hard cheese production process.

PO^2 ontology description:

1. Part concerning the production *process* that contains the concepts: process, itinerary and step. An *itinerary* is an execution of a production *process*, i.e., a set of interrelated steps. It allows for description of different possible technological processes. A *step* is characterised by its participants and its temporal duration.
2. Part concerning the *participant* that contains the concepts: product, mixture, material and method. A participant may be a *mixture*, a *material* or a *method*. Each participant is characterised by its experimental conditions. A *mixture* is characterised by its composition (i.e. the *products* it contains). For instance, the *mixture* "cheese" is composed of several *products* such as "proteins" and "lipids". Fig. 2 gives an excerpt of the milk products hierarchy in the PO^2 domain ontology. Notice that the concept "hard cheese" belongs to the concept "cheese", which belongs to the concept "milk products" which belongs and to the core concept "product".
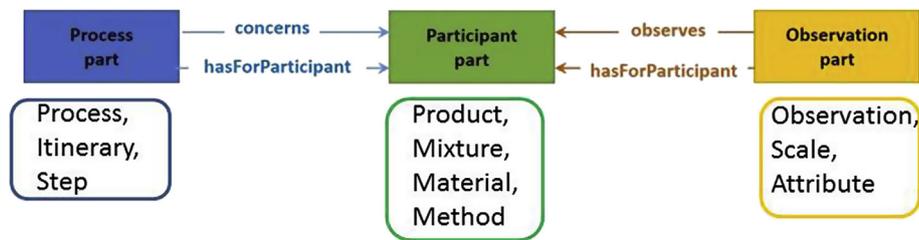
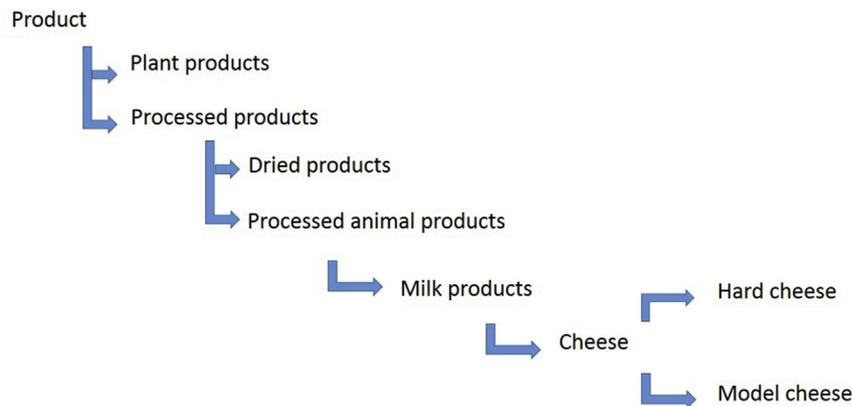**Fig. 1.** The three main parts of the PO^2 core ontology.

Fig. 2. An excerpt of the milk products hierarchy in the PO^2 domain ontology.

3. Part concerning the *observation* that contains the concepts: observation, scale and attribute, which can be symbolic or quantitative. An *observation* observes a *participant* (e.g., a mixture, a material) at a given *scale* during a given *step*. An observation stored in BaGaTel allows one to have both raw and computed data about several different *attributes* such as pH, taste intensity or energy consumption of the equipment.

PO^2 ontology is implemented in OWL 2 (https://www.w3.org/2001/sw/wiki/OWL) and its core layer is available on the Agro-Portal repository (http://agroportal.lirmm.fr/ontologies/PO2) under the license Creative Commons Attribution International 4.0 International (CC BY 4.0) (https://creativecommons.org/licenses/by/4.0/).

PO^2 ontology guided the structure of the BaGaTel database, as detailed in Supplementary material file SF5, integrating and storing all the available data presented in Section 2.1.1. These data are represented with the standard format Resource Description Framework (RDF). RDF integrated data can be queried by using semantic queries expressed in SPARQL.

2.3. Implementation of data

In a first step, experts in food science, especially in food process, oral physiology, sensory perception and eco-design worked to develop a consensual vocabulary gathered in PO^2 ontology, which has been used for the structuring of BaGaTel.

In a second step, data were utilised in BaGaTel from different collaborative national or international projects, theses and scientific publications from scientists involved in the NutriSensAl project (http://plasticnet.grignon.inra.fr/PortailNutriSensAl/). BaGaTel contains data for different types of parameters covering different steps of the process, raw materials (useful data to perform LCA), composition of the food product at the different steps, structure of the food product, evaluation of the environmental, sensory and nutritional qualities of the final product, taking into account food breakdown in the mouth and digestive track when relevant.

In a third step, data from different projects dealing with the same type of food product were compared. The challenge was then to identify data that were comparable. For example, concerning cheese products, measurements made using various devices and different calibrations of the device, sensory evaluations were made with various panels and on products from different sensory spaces, according to the specific project.

Thanks to our network of scientific and technical experts in the field, it was possible to identify types of cheeses and parameters that could be compared among them. After selection of the most relevant parameters, with the aim of always having the same parameter and same unit of measurement, some transformations of data and adjustments were done. For example, all scores obtained in sensory analysis were standardised on a 10-point scale. Physico-chemical data of composition were all expressed as gram of the item concerned (lipid, protein, dry matter…) per kilogram of total product.

All these modifications and adjustments must absolutely be validated by experts in food science and by scientists in charge of the thesis, the publication or the project.

3. Results and discussion

Results presented in this section aim at demonstrating the added value of the PO^2 ontology for querying the BaGaTel database to answer questions that need data from different projects. The integrated use of data coming from different projects allows for estimation of missing data in one specific project from a combination of data from other projects.

3.1. Combining data about hard cheese composition, process, sensory quality and rheological parameters

We chose to focus on "hard cheeses"; these cheeses have common characteristics, such as composition, technological process and structure. In the BaGaTel database two projects, PRASEL and TRUEFOOD, on real hard cheeses had common steps in their

processing: renneting, draining, pressing, moulding, demoulding, and ripening.

By using the RDF integrated data, BaGaTel was queried with Q1 (Which hard cheese sample has the highest value for "taste intensity"?) by a SPARQL query shown in Supplementary material Fig. S1. The answer, presented in Supplementary material Fig. S2, is sample "Fd24" that has a taste intensity of 6.82/10. This sample also had the highest aroma intensity (6.45/10). This sample belongs to TRUEFOOD project.

The answer to Q2 ("What is the composition of this sample at different steps in the process?") is given in Table 2. The lipid and protein content are given at three steps "milk reception", "in the vat" and "cheese ripening". Table 2 also shows the answer given by BaGaTel to Q3 ("What are the characteristics measured on this sample?"). The pH value has been measured at different steps ("moulding", "step in the vat", "demoulding") and the sensory properties have been measured at the "chewing" step.

No rheological properties could be found in BaGaTel database to answer to Q3, while these parameters are important to evaluate the quality of cheeses. Hence, it was necessary to estimate this missing data using data from other projects imported into BaGaTel.

The rheological parameters such as Young modulus have been demonstrated to be relevant to predict masticatory activity, especially with the aim of formulating food products for specific populations such as elderly people (Engelen, Fontijn-Tekamp, & Bilt, 2005; Peyron, Maskawi, Woda, Tanguay, & Lund, 1997). They can be used to select samples that are easily broken down in the mouth. Samples for this population are known to present reduced masticatory facilities. The aim of the following query is to estimate rheological parameters for the sample Fd24 using relationships between rheological parameters and sensorial parameters in BaGaTel.

We focused on one specific rheological parameter, Young modulus. The literature shows a good correlation between the Young modulus and the sensorial parameter "firmness" for cheese products (Foegeding et al., 2003), and for other food products like banana (Finney, Ben-Gera, & Massie, 1967) and apple (Mohsenin, Cooper, Hammerle, Fletcher, & Tukey, 1965).

By using this correlation, we wanted to estimate the Young modulus value for sample Fd24.

Fig. 3 contains for all cheese samples in BaGaTel the value for the sensory characteristic "firmness" as a function of the Young modulus in kPa. However, a poor correlation is observed between "firmness" and "Young modulus" if we consider all the cheese samples in BaGaTel ($R^2 = 0.37$).

However, if we focus only on cheese samples belonging to "hard cheese" category, the linear relationship between "firmness" and "Young modulus" (Fig. 3, bottom right), is higher ($R^2 = 0.74$) [Equation (1)]:

$$Firmness = 0.0086 \times Young\ modulus + 1.95\ (R^2 = 0.74) \quad (1)$$

To estimate Young modulus' of the sample Fd24, a selection was made among all samples. The database was queried to identify samples characterised by the following criteria (query denoted Q3bis):

1. belongs to the category "hard cheese" in the ontology,
2. has been characterised for its "firmness" and for its "Young modulus",
3. has a composition in terms of lipid content and protein content (in mg kg^{-1} of cheese) at the "ripening" step of the same order of magnitude as the sample Fd24 ($\pm$20%).

The assumption in criterion 3 is that the protein network and the lipid content play an important role in the firmness of a product (Fu et al., 2018). Therefore, the query Q3bis imposed that the protein and lipid content values of a sample could not be lower or higher than 20% of the respective values for the sample Fd24.

Only six samples matched all the criteria of the query Q3bis. A linear relationship exists between "firmness" and "Young modulus" (Equation (2)):

$$Firmness = 0.0074 \times Young\ Modulus + 3.1899\ (R^2 = 0.83) \quad (2)$$

The correlation between "firmness" and "Young modulus" is higher considering a selected set of samples [Equation (2), $R^2 = 0.83$, Fig. 4] than for samples matching only the first two criteria of Q3bis [Equation (1), $R^2 = 0.74$, Fig. 3].
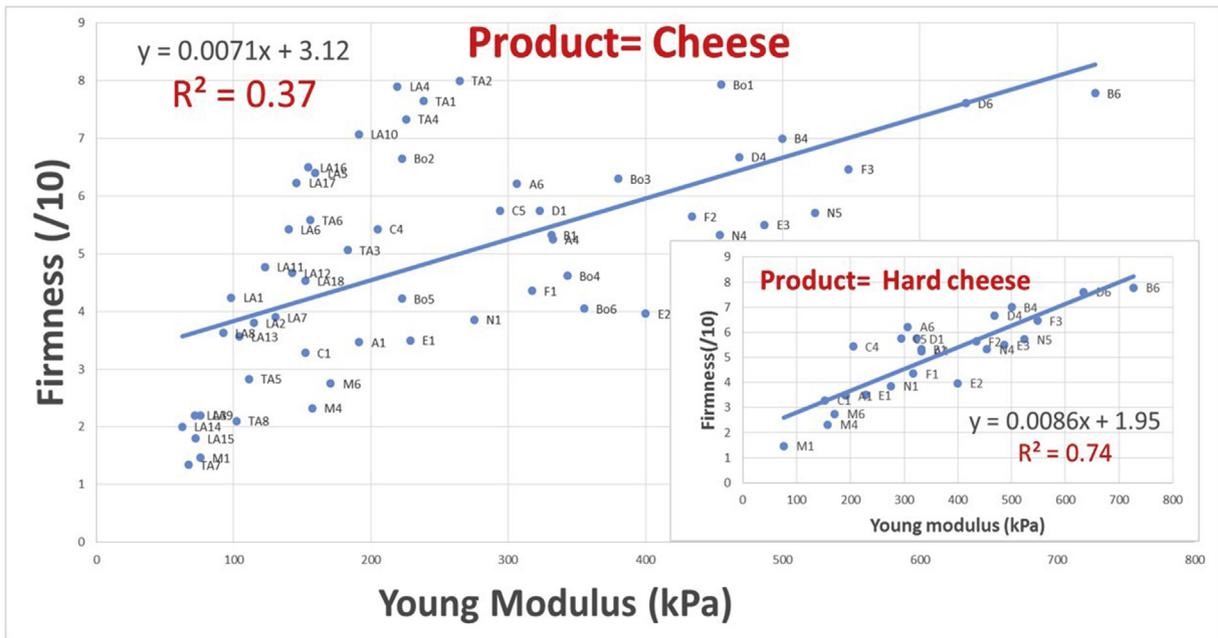
Lipid and protein contents, firmness characteristic and Young modulus parameters of the six samples found in the BaGaTel database to answer the criteria from Q3bis are presented in Supplementary material Fig. S3. As given in Table 2, the "firmness" value for the sample Fd24 was 4.64. Among the six samples from Fig. 4, one has a "firmness" value lower than 4.64 while the other five samples have "firmness" values higher than 4.64. We chose the sample with the lower value (i.e., sample C1) and the sample with the value just above (i.e., sample C4). Therefore, the value of the Young modulus of the sample Fd24 was estimated between the values of the Young modulus of sample C1, i.e., 152.37 kPa, as lower limit and the value of the Young modulus of sample C4, i.e., 205.13 kPa, as upper limit (Fig. 4).

To have another estimation of the Young modulus value of sample Fd24, it was also possible to use the linear relationships between firmness and Young modulus for hard cheeses. As sample Fd24 has a firmness rating of 4.64, the calculated estimation of the value for the Young modulus using Equation (2) was 196 kPa. This value is in the first estimated interval [152.37 kPa; 205.13 kPa]. An estimation with Equation (1) that only takes into account criteria 1 and 2 of query Q3bis gave an estimation of 312.5 kPa for Young modulus value for sample Fd24. This value is out of the previous estimated interval. This validates the usefulness of criterion 3 in Q3bis for this estimation, assuming that the relationship between firmness and Young modulus depends on the protein and lipid content of the cheese.

Therefore, this approach can estimate the missing rheological parameters for hard cheese samples by combining and reusing data from different projects. This estimation was only obtained by mathematical thinking and has to be confirmed by a computation of numerous data sets or by performing real experiments.

3.2. Combining data about hard cheese composition, processing and environmental impact

The other questions, Q4 and Q5, address environmental concerns to BaGaTel database. No environmental data have been collected for the Fd24 sample. Therefore, this section will show how

Table 2
Characteristics measured on the sample Fd24 at different steps of the production process.

Step	Characteristic	Value	Unit
Milk reception	Protein	35.0	g L^{-1}
Milk reception	Lipid	44.4	g L^{-1}
Step in the vat	Protein	35.1	g L^{-1}
Step in the vat	Lipid	38.5	g L^{-1}
Step in the vat	pH	6.61	None
Moulding	pH	6.32	None
Demoulding	pH	5.07	None
Chewing	Salty	4.55	/10
Chewing	Fat	8.27	/10
Chewing	Firmness	4.64	/10
Chewing	Aroma intensity	6.45	/10
Chewing	Taste intensity	6.82	/10
Cheese ripening	Protein	313.4	g kg^{-1}
Cheese ripening	Lipid	344.74	g kg^{-1}

Fig. 3. Correlation between firmness value and Young modulus value for all cheeses samples in BaGaTel and correlation between firmness value and Young modulus value for "hard cheeses" samples (Prasel project).
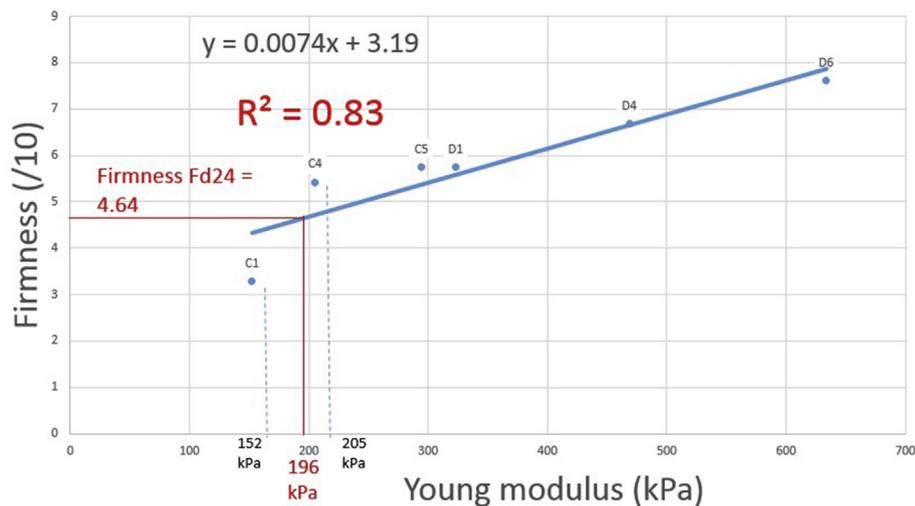
Fig. 4. Correlation between firmness value and Young modulus value for hard cheeses samples (Prasel project) answering to the criteria Q3bis.

BaGaTel database could be used to estimate environmental impact of sample Fd24.

Environmental impact has been quantified by using LCA in the CellExtraDry project (2014—2016), in the case of production of stabilised micro-organisms. BaGaTel database can be very useful to support life cycle inventory of data [step (ii) in LCA methodology] since it already contains data coming from LCA inventories and results from the CellExtraDry project. The answer to the question Q4 ("Which parameters need to be measured throughout the process to perform a Life Cycle Assessment of the sample?"), is summarised in Supplementary material Table S1. Measured parameters (e.g., electricity consumption, mass of water, volume of compressed air) are obtained as well as the step when data related to these parameters have been collected. That provides detailed items that have been previously measured to perform LCAs. LCA practitioner then has to consider if these are relevant for their system, and if other similar items have to be included. For instance, the mass of saccharose used for the fermentation step in CellExtraDry project might not be relevant to a cheese production system. However, by analogy, a LCA practitioner can easily extrapolate this result to conclude that it would be important to measure the initial mass of raw milk. Combination of BaGaTel querying and expertise of a LCA practitioner is crucial to state all relevant data that have to be gathered. This approach may help LCA practitioners to transfer knowledge from one domain to another domain by suggesting relevant parameters to be measured. This is possible thanks to the high-level description of the available data using the PO^2 ontology.

The BaGaTel database can therefore be also queried to obtain relevant quantified data on specified items. For instance, from Supplementary material Table S1, it can be deduced that energy consumption is an important parameter, since it was measured at different steps (fermentation, centrifugation, drying). The query Q5,

Table 3
Energy consumption at different steps for French hard cheese production.[a]

Step	Material	Power (kW)	Usage duration (h)	Electricity consumption (kWh)
Skimming	Cream separator	0.065	1	0.065
Cooling	Refrigerated tank	14.52	2	29.04
Step in the vat	Heating equipment	9	0.75	6.75
Step in the vat	Heating equipment	9	1.15	9.45
Cheese ripening	Maturing cellar 1	1.25	720	900
Cheese ripening	Maturing cellar 2	1.25	1080	1350
Cheese ripening	Maturing cellar 3	1.25	1800	2250

[a] Results come from TRUEFOOD (2006—2010) project.

("What are the values of the relevant parameters needed to conduct the Life Cycle Assessment?"), could be reformulated as "What is the energy consumption of a French hard cheese production process?". Unfortunately, no data were directly available in the BaGaTel database to answer this question. Nevertheless, because PO[2] ontology also integrates material and methods related to the experiments, the BaGaTel database could be queried on this to evaluate the energy consumption during the different steps of hard cheese production presented in Section 2.1.2. For each step, the material used could be identified, duration of its use and power of the equipment. Then energy consumption was calculated by using Equation (3):

$$\text{Energy consumption} = \text{Power of equipment} \times \text{Use duration of equipment} \tag{3}$$

Table 4
Mass flow data involved in the production of sample Fd24.[a]

| Step	Material	Amount
Milk reception	Milk	125 L
Vat cleaning	Sodium metasilicate	200 g
	Sodium carbonate	200 g
	Phosphoric acid	250 g
	Peracetic acid	250 g
	Hydrogen peroxide	250 g
	Ethanoic acid	250 g
Pickling	Pickle	25 L
Cheese Ripening 1	Pickle and Smear	20 g
Cheese Ripening 2	Pickle and Smear	20 g

[a] Results come from TRUEFOOD (2006—2010) project.

Results are presented in Table 3.

Mass flow data concerning sample Fd24 could also be obtained by querying BaGaTel database, these are presented in Table 4.

Finally, using data described in Tables 3 and 4, supplemented by data and emission factors of Ecoinvent database, a simplified LCA could be performed. This simplified LCA referred to the production of one cheese of 10.5 kg ripened over 120 days, corresponding to the use of 115 L of raw milk. A simplified LCA was performed, but not detailed, because this paper aims at showing how the BaGaTel database can be helpful in LCA studies, but does not claim to provide thorough LCA results. For this reason, LCA results will not be deeply discussed. Results for climate change, ozone depletion and acidification indicators are presented in Fig. 5. For clarity reasons, other indicators are not shown. Milk reception was the main contributor to climate change and acidification, as well as for ten other indicators (not shown), due to raw milk production. On the other side, cheese ripening was the main contributor to ozone depletion, as well as for three other indicators (not shown), due to energy demand by the ripening rooms. These results are consistent with other LCA studies which showed that raw milk production and energy use during ripening were the main causes of environmental impacts of ripened cheese (Berlin, 2002; Finnegan, Yan, Holden, & Goggins, 2017).

Finally, when a LCA practitioner has finished: LCA inventory data and LCA results can be stored in the BaGaTel database, together with all the corresponding metadata (e.g. system boundaries, functional unit, impact assessment method) necessary to eventually re-use them. From these inventory and results, LCA data
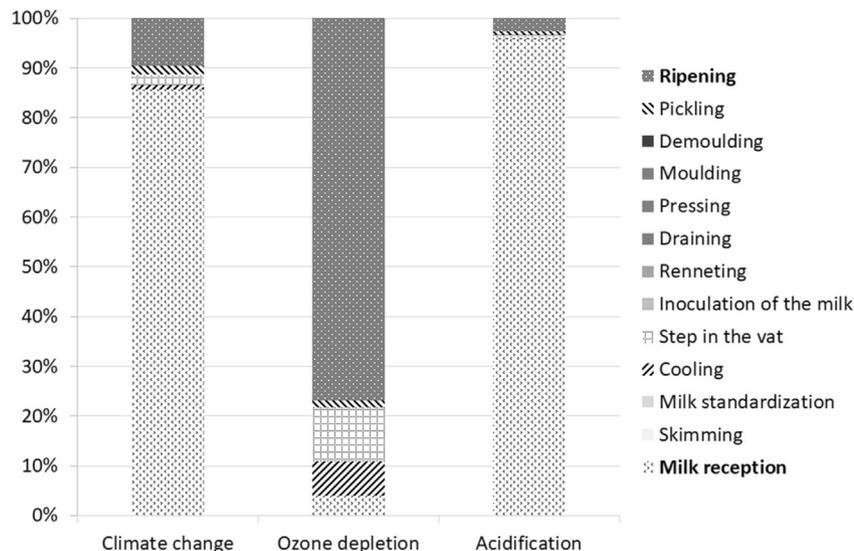
Fig. 5. Climate change, Ozone depletion and Acidification indicators of the contribution analysis for the Life Cycle Assessment of the cheese produced to obtain sample Fd24 — ILCD midpoint method.

could enrich existing databases (e.g., LCA Food database, World Food LCA database, EcoInvent database) that are currently poorly documented with regards to impacts due to processes as a function of (i) the food product submitted to the process and (ii) the process operating conditions.

4. Conclusions

This study demonstrates that the common vocabulary and structure provided by PO^2 ontology allow for combining and integrating data from different projects in a common database, BaGaTel, and therefore giving relevant answers to several questions. BaGaTel can add value to evaluate the relevance of a data set to be used for the estimation of missing data for a specific sample, taking into account the data available. On the question addressing the formulation of hard cheese, BaGaTel database could be useful in estimate missing data on cheese rheology, by restricting the query to a set of samples with similar measured characteristics. On the question about eco-design of a food product, it gave hints about relevant data that have to be acquired and could also be useful in estimate missing data. Thanks to the fact that data on cheese quality, process and eco-design are in the same database, it will now be possible to estimate missing data on the environmental impact of projects only focused on food quality and also estimate relevant parameters for food quality in samples from project only focused on eco-design. Such an approach is very useful for knowledge and data capitalisation, as well as to produce new knowledge and data by combining and integrating existing resources. As a consequence, the number of available observations in the database is a key-parameter for making BaGaTel structured by PO^2 valuable. At the same time, such a tool makes it possible to identify the lack of data, which have to be estimated or experimentally measured, to be able to answer challenging questions.

Further work should be done to investigate to what extent open data available on the LOD (Linked Open Data) could be reused in connection with PO^2. In addition, by combining the presented tools with artificial intelligence techniques, a decision support tool could be developed to formulate future foods answering specific quality properties and produced with a controlled environmental impact. This would be a big step towards more sustainable food systems.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.idairyj.2019.01.003.

References

AFNOR. (2009). *Analyse sensorielle, Recueil de normes. Saint Denis la Plaine*. France: AFNOR.

Berlin, J. (2002). Environmental life cycle assessment (LCA) of Swedish semi-hard cheese. *International Dairy Journal, 12*, 939—953.

Bigaski Ribeiro, J. C., Granato, D., Masson, M. L., Andriot, I., Mosca, A. C., Salles, C., et al. (2016). Effect of lactobionic acid on the acidification, rheological properties and aroma release of dairy gels. *Food Chemistry, 207*, 101—106.

Bizer, C. (2013). Interlinking scientific data on a global scale. *Data Science Journal, 12*, GRDI6—GRDI12.

Boisard, L., Andriot, I., Arnould, C., Achilleos, C., Salles, C., & Guichard, E. (2013). Structure and composition of model cheeses influence sodium NMR mobility, kinetics of sodium release and sodium partition coefficients. *Food Chemistry, 136*, 1070—1077.

Boisard, L., Andriot, I., Martin, C., Septier, C., Boissard, V., Salles, C., et al. (2014). The salt and lipid composition of model cheeses modifies in-mouth flavour release and perception related to the free sodium ion content. *Food Chemistry, 145*, 437—444.

Dibie, J., Dervaux, S., Doriot, E., Ibanescu, L., & Pénicaud, C. (2016). [MS]^{2}O—A multi-scale and multi-step ontology for transformation processes: Application to micro-organisms. In O. Haemmerlé, G. Stapleton, & C. Faron Zucker (Eds.), *Graph-based representation and reasoning. Proceedings of the 22nd international conference on conceptual structures* (pp. 163—176). Berlin, Germany: Springer.

Doan, A., Halevy, A. Y., & Ives, Z. G. (2012). *Principles of data integration*. Waltham, MA, USA: Morgan Kaufmann.

Doyennette, M., Deleris, I., Saint-Eve, A., Gasiglia, A., Souchon, I., & Trelea, I. C. (2011). The dynamics of aroma compound transfer properties in cheeses during simulated eating conditions. *Food Research International, 44*, 3174—3181.

Engelen, L., Fontijn-Tekamp, A., & van der Bilt, A. (2005). The influence of product and oral characteristics on swallowing. *Archives of Oral Biology, 50*, 739—746.

European Commission. (2011). *International reference life cycle data system (ILCD) handbook: General guide for life cycle assessment: Provisions and action steps (Institute for environment and sustainability)*. Luxembourg: Publications Office.

Feron, G., Ayed, C., Qannari, E. M., Courcoux, P., Laboure, H., & Guichard, E. (2014). Understanding aroma release from model cheeses by a statistical multiblock approach on oral processing. *PLoS One, 9*(4), e93113.

Finnegan, W., Yan, M., Holden, N. M., & Goggins, J. (2017). A review of environmental life cycle assessment studies examining cheese production. *International Journal of Life Cycle Assessment, 23*, 1773—1787.

Finney, E. E., Ben-Gera, I., & Massie, D. R. (1967). An objective evaluation of changes in firmness of ripening bananas using a sonic technique. *Journal of Food Science, 32*, 642—646.

Foegeding, E. A., Brown, J., Drake, M., & Daubert, C. R. (2003). Sensory and mechanical aspects of cheese texture. *International Dairy Journal, 13*, 585—591.

Fox, P. F. (2011). Cheese — Overview. In J. W. Fuquay (Ed.), *Encyclopedia of dairy sciences* (2nd ed., pp. 533—543). San Diego, CA, USA: Academic Press.

Fu, W., Watanabe, Y., Inoue, K., Moriguchi, N., Fusa, K., Yanagisawa, Y., et al. (2018). Effects of pre-cooked cheeses of different emulsifying conditions on mechanical properties and microstructure of processed cheese. *Food Chemistry, 245*, 47—52.

Gierczynski, I., Labouré, H., Sémon, E., & Guichard, E. (2007). Impact of hardness of model fresh cheese on aroma release: In vivo and in vitro study. *Journal of Agricultural and Food Chemistry, 55*, 3066—3073.

Guarino, N., Oberle, D., & Staab, S. (2009). What is an ontology? In S. Staab, & R. Studer (Eds.), *Handbook on ontologies* (pp. 1—17). Berlin, Germany: Springer.

Guichard, E., Repoux, M., Qannari, E. M., Laboure, H., & Feron, G. (2017). Model cheese aroma perception is explained not only by in vivo aroma release but also by salivary composition and oral processing parameters. *Food and Function, 8*, 615—628.

Hailu, Y., Hansen, E. B., Seifu, E., Eshetu, M., Petersen, M. A., Lametsch, R., et al. (2018). Rheological and sensory properties and aroma compounds formed during ripening of soft brined cheese made from camel milk. *International Dairy Journal, 81*, 122—130.

Ibanescu, L., Dibie, J., Dervaux, S., Guichard, E., & Raad, J. (2016). PO^2 — A process and observation ontology in food science. Application to dairy gels. In E. Garoufallou, I. Subirats Coll, A. Stellato, & J. Greenberg (Eds.), *Metadata and semantics research (pp. 155—165; proceedings of 10th International Conference, MTSR 2016. Göttingen, Germany)*. Berlin, Germany: Springer.

Igos, E., Benetto, E., Baudin, I., Tiruta-Barna, L., Mery, Y., & Arbault, D. (2013). Cost-performance indicator for comparative environmental assessment of water treatment plants. *Science of the Total Environment, 443*, 367—374.

ISO. (2006). *Environmental management - life cycle assessment — Principles and framework*. ISO 14040:2006 (2nd ed.). Geneva, Switzerland: International Organisation for Standardisation.

Lauverjat, C., Deleris, I., Trelea, I. C., Salles, C., & Souchon, I. (2009). Salt and aroma compound release in model cheeses in relation to their mobility. *Journal of Agricultural and Food Chemistry, 57*, 9878—9887.

Lawrence, G., Buchin, S., Achilleos, C., Berodier, F., Septier, C., Courcoux, P., et al. (2012). In vivo sodium release and saltiness perception in solid lipoprotein matrices. 1. Effect of composition and texture. *Journal of Agricultural and Food Chemistry, 60*, 5287—5298.

Lawrence, G., Salles, C., Palicki, O., Septier, C., Busch, J., & Thomas-Danguin, T. (2011). Using cross-modal interactions to counterbalance salt reduction in solid foods. *International Dairy Journal, 21*, 103—110.

Lousteau-Cazalet, C., Barakat, A., Belaud, J.-P., Buche, P., Busset, G., Charnomordic, B., et al. (2016). A decision support system for eco-efficient biorefinery process comparison using a semantic approach. *Computers and Electronics in Agriculture, 127*, 351—363.

Mohsenin, N. N., Cooper, H. E., Hammerle, J. R., Fletcher, S. W., & Tukey, L. D. (1965). "Readiness for harvest" of apples as affected by physical and mechanical properties of the fruit. *University Park*. PA, USA: Pennsylvania State University, College of Agriculture, Agricultural Experiment Station.

Mosca, A. C., Andriot, I., Guichard, E., & Salles, C. (2015). Binding of Na+ ions to proteins: Effect on taste perception. *Food Hydrocolloids, 51*, 33—40.

Muljarto, A. R., Salmon, J.-M., Charnomordic, B., Buche, P., Tireau, A., & Neveu, P. (2017). A generic ontological network for Agri-food experiment integration — Application to viticulture and winemaking. *Computers and Electronics in Agriculture, 140*, 433—442.

Muljarto, A.-R., Salmon, J.-M., Neveu, P., Charnomordic, B., & Buche, P. (2014). Ontology-based model for food transformation processes — Application to

winemaking. In S. Closs, R. Studer, E. Garoufallou, & M. A. Sicilia (Eds.), *Metadata and semantics research* (pp. 329—343; *proceedings of 8th International Conference, MTSR 2014. Karlsruhe, Germany*). Berlin, Germany: Springer.

Noël, Y., Zannoni, M., & Hunter, E. A. (1996). Texture of Parmigiano Reggiano cheese: Statistical relationships between rheological and sensory variates. *Lait, 76*, 243—254.

Pénicaud, C., Monclus, V., Perret, B., Passot, S., & Fonseca, F. (2018). Life cycle assessment of the production of stabilized lactic acid bacteria for the environmentally-friendly preservation of living cells. *Journal of Cleaner Production, 184*, 847—858.

Peyron, M. A., Maskawi, K., Woda, A., Tanguay, R., & Lund, J. P. (1997). Effects of food texture and sample thickness on mandibular movement and hardness assessment during biting in man. *Journal of Dental Research, 76*, 789—795.

Phan, V. A., Yven, C., Lawrence, G., Chabanet, C., Reparet, J. M., & Salles, C. (2008). In vivo sodium release related to salty perception during eating model cheeses of different textures. *International Dairy Journal, 18*, 956—963.

Pizzuti, T., Mirabelli, G., Grasso, G., & Paldino, G. (2017). MESCO (MEat Supply Chain Ontology): An ontology for supporting traceability in the meat supply chain. *Food Control, 72*, 123—133.

Pizzuti, T., Mirabelli, G., Sanz-Bobi, M. A., & Gomez-Gonzalez, F. (2014). Food Track & Trace ontology for helping the food traceability control. *Journal of Food Engineering, 120*, 17—30.

Roy, P., Nei, D., Orikasa, T., Xu, Q. Y., Okadome, H., Nakamura, N., et al. (2009). A review of life cycle assessment (LCA) on some food products. *Journal of Food Engineering, 90*, 1—10.

Suárez-Figueroa, M. del C., Gómez-Pérez, A., & Fernández-López, M. (2012). The NeOn methodology for ontology engineering. In M. C. Suárez-Figueroa, A. Gómez-Pérez, E. Motta, & A. Gangemi (Eds.), *Ontology engineering in a networked world* (pp. 9—34). Berlin, Germany: Springer.

Tarrega, A., Yven, C., Sémon, E., & Salles, C. (2008). Aroma release and chewing activity during eating different model cheeses. *International Dairy Journal, 18*, 849—857.

Tarrega, A., Yven, C., Sémon, E., & Salles, C. (2011). In-mouth aroma compound release during cheese consumption: Relationship with food bolus formation. *International Dairy Journal, 21*, 358—364.

Tukker, A., Huppes, G., Guinée, J., Heijungs, R., de Koning, A., van Oers, L., et al. (2006). *Environmental impact of products (EIPRO). Analysis of the life cycle environmental impacts related to the final consumption of the EU-25*. Seville, Spain: European Commission, Joint Research Centre.

Weidema, B. P., Wesnaes, M., Hermansen, J., Kristensen, T., Halberg, N., Eder, P., et al. (2008). *Environmental improvement potentials of meat and dairy products (IMPRO)* (p. 196). Seville, Spain: European Commission, Joint Research Centre.