



# Food Concept Alignment KG



Patrice Buche, Liliana Ibanescu  
27 juin 2022

Atelier ROCED @ PFIA 2022

## Contexte

- Un FUI impliquant des partenaires industriels
- Objectif: évaluer le score nutritionnel d'une recette

Valeur nutritive	
par 175 g	
Teneur	% valeur quotidienne
<b>Calories</b> 130	
<b>Lipides</b> 0,5 g	<b>1 %</b>
saturés 0,3 g + trans 0 g	<b>2 %</b>
<b>Cholestérol</b> 4 mg	
<b>Sodium</b> 125 mg	<b>5 %</b>
<b>Glucides</b> 26 g	<b>8 %</b>
Fibres 0 g	<b>0 %</b>
Sucres 26 g	
<b>Protéines</b> 8 g	
Vitamine A 8 %	Vitamine C 4 %
Calcium 25 %	Fer 0 %

<https://www.extenso.org/article/votre-guide-nutritionnel-a-l-epicerie/>

## Contexte

- Un FUI impliquant des partenaires industriels
- Objectif: évaluer le score nutritionnel d'une recette
  1. Trouver les valeurs nutritionnelles des aliments dans des tables de composition

riboflavin	0.008000	milligram	per 100g edible portion	best estimate	imputed/estimated, generic	Analytical methods	File or Database
thiamin	0.100000	milligram	per 100g edible portion	best estimate	imputed/estimated, generic	Analytical methods	File or Database
vitamin C (ascorbic acid)	6.500000	milligram	per 100g edible portion	best estimate	imputed/estimated, generic	Analytical methods	File or Database
lycopene	0.000000	microgram	per 100g edible portion	best estimate	Analytical result(s)	Analytical methods	Article in Journal
beta-cryptoxanthin	0.000000	microgram	per 100g edible portion	best estimate	Analytical result(s)	Analytical methods	Article in Journal

## Contexte

- Un FUI impliquant des partenaires industriels
- Objectif: évaluer le score nutritionnel d'une recette
  1. Trouver les valeurs nutritionnelles des aliments dans des tables de composition
  2. Calculer le score de la recette
- Défi: les tables de composition des aliments (TCA)
  - par pays
  - des nomenclatures des aliments différentes
  - incomplètes
  - payantes

## Notre cas d'usage

Comment évaluer des valeurs manquantes (Fe, vitamine B12) d'un aliment dans une TCA?

➤ Retrouver ces valeurs dans une **autre TCA** pour un **aliment similaire**

### Mise en œuvre:

- Tables de composition utilisées:
  - CIQUAL de ANSES
  - SIREN de USDA
- Un algo qui trouve le plus proche aliment et la sa famille dans FoodON
- Un outil Multi-DB explorer

## LanguaL (<http://langual.org/>)

- Thésaurus multilingue pour la classification multi-facettes des aliments
  - Un aliment peut être décrit par une combinaison des descripteurs
  - Les descripteurs sont standardisés et hiérarchisés

Table 2. Languag descriptors for food **Cooked pork shoulder**, choice used in Ciqal database

LANGUAL CODE	LANGUAL TERM
A0279	CURED MEAT (US CFR)
A0797	PRESERVED MEAT (EUROFIR)
B1136	SWINE
C0270	SKELETAL MEAT PART, WITHOUT BONE AND SKIN, WITHOUT SEPARABLE FAT
E0137	SLICED
F0014	FULLY HEAT-TREATED
G0003	COOKING METHOD NOT APPLICABLE
H0253	CURED OR AGED
H0367	SALT ADDED
J0100	PRESERVED BY ADDING CHEMICALS
J0131	PRESERVED BY CHILLING
P0024	HUMAN CONSUMER, NO AGE SPECIFICATION
Z0010	CHOICE GRAD
Z0043	SHOULDER (MEAT CUT)

Table 2. Languag descriptors for food Cooked pork shoulder, choice used in Ciqal database

LANGUAL CODE	LANGUAL TERM
A0279	CURED MEAT (US CFR)
A0797	PRESERVED MEAT (EUROFIR)
B1136	SWINE
C0270	SKELETAL MEAT SEPARABLE FAT
E0137	SLICED
F0014	FULLY HEAT-TREATED
G0003	COOKING METHOD
H0253	CURED OR AGE
H0367	SALT ADDED
J0100	PRESERVED BY /
J0131	PRESERVED BY C
P0024	HUMAN CONSUM
Z0010	CHOICE GRAD
Z0043	SHOULDER (MEAT)

- A. PRODUCT TYPE
  - DIETARY SUPPLEMENT
  - FOOD ADDITIVES
  - PRODUCT TYPE, EUROPEAN UNION
    - CIAA FOOD CLASSIFICATION FOR FOOD ADDITIVES
    - CLASSIFICATION OF PRODUCTS OF PLANT AND ANIMAL ORIGIN, EUROPEAN COMMUNITY
    - EFSA FOOD CLASSIFICATION AND DESCRIPTION SYSTEM FOR EXPOSURE ASSESSMENT (EFSA FOODEX2)
    - EUROCODE 2 FOOD CLASSIFICATION
    - EUROFIR FOOD CLASSIFICATION
      - BEVERAGE (NON-MILK) (EUROFIR)
      - EGG OR EGG PRODUCT (EUROFIR)
      - FAT OR OIL (EUROFIR)
      - FRUIT OR FRUIT PRODUCT (EUROFIR)
      - GRAIN OR GRAIN PRODUCT (EUROFIR)
      - MEAT OR MEAT PRODUCT (EUROFIR)
        - MEAT ANALOGUE (EUROFIR)
        - MEAT DISH (EUROFIR)
        - OFFAL (EUROFIR)
        - POULTRY MEAT (EUROFIR)
        - PRESERVED MEAT (EUROFIR)
        - RED MEAT (EUROFIR)
        - SAUSAGE OR SIMILAR MEAT PRODUCT (EUROFIR)
      - MILK, MILK PRODUCT OR MILK SUBSTITUTE (EUROFIR)
      - MISCELLANEOUS FOOD PRODUCT (EUROFIR)
      - NUT, SEED OR KERNEL (EUROFIR)
      - PRODUCT FOR SPECIAL NUTRITIONAL USE OR DIETARY SUPPLEMENT (EUROFIR)
      - SEAFOOD OR RELATED PRODUCT (EUROFIR)
      - SUGAR OR SUGAR PRODUCT (EUROFIR)
      - VEGETABLE OR VEGETABLE PRODUCT (EUROFIR)
  - EUROPEAN FOOD GROUPS (EFG)
  - PRODUCT TYPE, INTERNATIONAL





## LanguaL (<http://langual.org/>)

- Thésaurus multilingue pour la classification multi-facettes des aliments
  - Un aliment peut être décrit par une combinaison des descripteurs
  - Les descripteurs sont standardisés et hiérarchisés
- Des TCA sont indexés avec LanguaL
  - [http://langual.org/langual\\_indexed\\_datasets.asp](http://langual.org/langual_indexed_datasets.asp)
  - CIQUAL
  - SIREN
- [FoodON](#) reprend LanguaL

- La table de composition des aliments gérée par ANSES
- Sa version de 2017
  - Contient 2807 aliments
    - Dont 1797 (64%) décrits avec des facettes LanguaL (FoodOn)

FOODID	ORIGFDNM	ENGFDM	LANGUALCODES
1000	Pastis	Pastis (anise-flavoured spirit)	A0200 A0850 B1347 C0174 E0123 F0014 G0003 H0136 H0150 H0229 H0232 H0270 J0104 K0003 M0214 N0
1001	Eau de vie	Clear fruit brandy or eau-de-vie	A0200 A0850 B1140 C0167 E0123 F0014 G0003 H0232 H0253 H0270 J0003 K0003 M013
1002	Gin	Gin	A0277 A0850 B1324 C0133 E0123 F0014 G0003 H0229 H0232 H0270 J0003 K0003 M0214 N0001 P0024 R0001 Z0001
1003	Liqueur	Liqueur	A0240 A0850 B1347 C0174 E0123 F0014 G0003 H0136 H0227 H0232 H0270 J0003 K0003 M0214 N0001 P0188 R0001 Z0001
1004	Rhum	Rum	A0200 A0850 B1378 C0207 E0123 F0014 G0003 H0232 H0270 J0001 K0003 M0214 N0001 P0188 R0001 Z0001
1005	Whisky	Whisky	A0200 A0850 B1324 C0133 E0123 F0014 G0003 H0232 H0270 H0294 J0003 K0003 M0214 P0188 R0001 Z0001
1006	Vin doux	"Wine, white, sweet"	A0224 A0849 B1329 C0167 E0123 F0003 G0003 H0232 H0332 J0104 K0003 M0214 N0001 P0188 R0001 Z0
1007	Apéritif à base de vin ou vermouth	Wine-based aperitif	A0224 A0851 B2121 C0167 E0123 F0003 G0003 H0232 H0253 H0332 J0104 K0
1008	Vodka	Vodka	A0120 A0850 B1347 C0133 E0123 F0014 G0003 H0232 H0270 J0003 K0003 M0214 N0001 P0188 R0001 Z0001
1011	Apéritif anisé sans alcool	Wine or fortified wine-based aperitif	A0295 A0846 B1296 C0155 E0123 F0022 G0003 H0148 H0229 J0135
1012	Cocktail à base de rhum	Rum-based cocktail	A0278 A0851 C0005 E0123 H0136 H0227 H0232 H0270 H0332
1013	Cocktail à base de whisky	Whiskey-based cocktail	A0851
1014	Alcool pur	Pure alcohol	A0120 A0850 A1284 B1347 C0174 E0123 F0014 G0003 H0232 H0270 J0003 K0003 M0001 N0001 P0188 R0001 Z011
1015	Marsala	Marsala wine	A0224 A0849 B1329 C0140 E0123 F0003 G0003 H0232 H0332 J0104 K0003 M0214 N0001 P0188 R0263 Z0255

# CIQUAL

- La table de composition des aliments gérée par ANSES
- Sa version de 2017
  - Contient 2807 aliments
  - Dont 1797 (64%) décrits avec des facettes LanguaL (FoodOn)

FOODID	ORIGFDNM	ENGFDM	LANGUALCODES
1000	Pastis	Pastis (anise-flavoured spirit)	A0200 A0850 B1347 C0174 E0123 F0014 G0003 H0136 H0150 H0229 H0232 H0270 J0104 K0003 M0214 N0
1001	Eau de vie	Clear fruit brandy or eau-de-vie	A0200 A0850 B1140 C0167 E0123 F0014 G0003 H0232 H0253 H0270 J0003 K0003 M013
1002	Gin	Gin	A0277 A0850 B1324 C0133 E0123 F0014 G0003 H0229 H0232 H0270 J0003 K0003 M0214 N0001 P0024 R0001 Z0001
1003	Liqueur	Liqueur	A0240 A0850 B1347 C0174 E0123 F0014 G0003 H0136 H0227 H0232 H0270 J0003 K0003 M0214 N0001 P0188 R0001 Z0001
1004	Rhum	Rum	A0200 A0850 B1378 C0207 E0123 F0014 G0003 H0232 H0270 J0001 K0003 M0214 N0001 P0188 R0001 Z0001
1005	Whisky	Whisky	A0200 A0850 B1324 C0133 E0123 F0014 G0003 H0232 H0270 H0294 J0003 K0003 M0214 P0188 R0001 Z0001
1006	Vin doux	"Wine, white, sweet"	A0224 A0849 B1329 C0167 E0123 F0003 G0003 H0232 H0332 J0104 K0003 M0214 N0001 P0188 R0001 Z0
1007	Apéritif à base de vin ou vermouth	Wine-based aperitif	A0224 A0851 B2121 C0167 E0123 F0003 G0003 H0232 H0253 H0332 J0104 K0
1008	Vodka	Vodka	A0120 A0850 B1347 C0133 E0123 F0014 G0003 H0232 H0270 J0003 K0003 M0214 N0001 P0188 R0001 Z0001
1011	Apéritif anisé sans alcool	Wine or fortified wine-based aperitif	A0295 A0846 B1296 C0155 E0123 F0022 G0003 H0148 H0229 J0135
1012	Cocktail à base de rhum	Rum-based cocktail	A0278 A0851 C0005 E0123 H0136 H0227 H0232 H0270 H0332
1013	Cocktail à base de whisky	Whiskey-based cocktail	A0851
1014	Alcool pur	Pure alcohol	A0120 A0850 A1284 B1347 C0174 E0123 F0014 G0003 H0232 H0270 J0003 K0003 M0001 N0001 P0188 R0001 Z011
1015	Marsala	Marsala wine	A0224 A0849 B1329 C0140 E0123 F0003 G0003 H0232 H0332 J0104 K0003 M0214 N0001 P0188 R0263 Z0255

- Une version en OWL disponible sur Data INRAE (<https://doi.org/10.15454/6CEYU3>)

# CIQUAL

- La table de composition de
- Sa version de 2017
  - Contient 2807 aliments
  - Dont 1797 (64%) décrits avec

FOODID	ORIGFDNM	ENGFDMAM	LANGUALCODES
1000	Pastis	Pastis (anise-flavoured spirit)	A0200 A0850
1001	Eau de vie	Clear fruit brandy or eau-de-vie	
1002	Gin	Gin	A0277 A0850 B1324 C0133 E0123 F0014
1003	Liqueur	Liqueur	A0240 A0850 B1347 C0174 E0123 F0014
1004	Rhum	Rum	A0200 A0850 B1378 C0207 E0123 F0014
1005	Whisky	Whisky	A0200 A0850 B1324 C0133 E0123 F0014
1006	Vin doux	"Wine, white, sweet"	A0224 A0849
1007	Apéritif à base de vin ou vermouth	Wine-based	
1008	Vodka	Vodka	A0120 A0850 B1347 C0133 E0123 F0014
1011	Apéritif anisé sans alcool	Wine or fortified wine	
1012	Cocktail à base de rhum	Rum-based cocktail	A0277
1013	Cocktail à base de whisky	Whiskey-based cocktail	
1014	Alcool pur	Pure alcohol	A0120 A0850 A1284 B1329
1015	Marsala	Marsala wine	A0224 A0849 B1329 C0140 E0123

Annotations: 28911 Description: 28911

Annotations +

**skos:prefLabel** [language: en] @ x o

Cooked pork shoulder, choice

---

**skos:hiddenLabel** [language: en] @ x o

cook pork shoulder choice

Equivalent To +

---

SubClass Of +

- has\_facet some FOODON\_03400279
- has\_facet some FOODON\_03400797
- has\_facet some FOODON\_03411136
- has\_facet some FOODON\_03420270
- has\_facet some FOODON\_03430137
- has\_facet some FOODON\_03440014
- has\_facet some FOODON\_03460253
- has\_facet some FOODON\_03460367
- has\_facet some FOODON\_03470100
- has\_facet some FOODON\_03470131
- has\_facet some FOODON\_03480003
- has\_facet some FOODON\_03500001
- has\_facet some FOODON\_03520001
- has\_facet some FOODON\_03530010
- has\_facet some FOODON\_03530043

- Une **version en OWL** disponible sur Data INRAE (<https://doi.org/10.15454/6CEYU3>)

## SIREN

- Scientific Information and Retrieval Exchange Network of the US Food and Drug administration
- Contient 8618 (??) aliments
  - Dont 5137 décrits avec des facettes LanguaL (FoodOn)

# SIREN

- Scientific Information and Retrieval Exchange Network of the US Food and Drug administration
- Contient 8618 (??) aliments
  - Dont 5137 décrits avec des facettes LanguaL (FoodOn)

The screenshot displays the FoodOn ontology interface for the concept FOODON\_03310969. The interface is divided into two main panels: Annotations and Description.

**Annotations:** The top panel shows two annotations for the concept:

- `skos:prefLabel` [language: en] with the value `pork shoulder (cooked, cured)`.
- `skos:hiddenLabel` [language: en] with the value `pork shoulder cook cure`.

**Description:** The right panel shows the concept's relationships:

- Equivalent To:** A plus sign (+) indicates that there are equivalent concepts.
- SubClass Of:** A plus sign (+) indicates that there are subclasses. The list shows the concept is a subclass of several other concepts, each with a yellow dot and the text `has_facet some` followed by a FoodOn ID:
  - FOODON\_03400279
  - FOODON\_03411136
  - FOODON\_03420125
  - FOODON\_03430151
  - FOODON\_03440014
  - FOODON\_03460253
  - FOODON\_03480001
  - FOODON\_03500001
  - FOODON\_03530043

# SIREN

- Scientific Information and Retrieval Exchange Network of the US Food and Drug administration
- Contient 8618 (??) aliments
  - Dont 5137 décrits avec des facettes

The screenshot displays a web interface for the FOODON\_03310969 ontology. It features two main panels: 'Annotations' and 'Description'. The 'Annotations' panel shows two entries: 'skos:prefLabel [language: en] pork shoulder (cooked, cured)@en' and 'skos:hiddenLabel [language: en] pork shoulder cook cure'. The 'Description' panel shows 'Description: FOODON\_03310969' and a list of 'SubClass Of' relationships, each with a yellow circle icon and the text 'has\_facet some FOODON\_XXXXXX'.

- Une version en OWL disponible sur Data INRAE (<https://doi.org/10.15454/5LLGVY>)

## Notre cas d'usage

Comment évaluer des valeurs manquantes (Fe, vitamine B12) d'un aliment dans une TCA?

➤ Retrouver ces valeurs dans une **autre TCA** pour un **aliment similaire**

### Mise en œuvre:

- Tables de composition utilisées:
  - **CIQUAL de ANSES**
  - **SIREN de USDA**
- Un **algo** qui trouve le plus proche aliment et la sa famille dans FoodON
- Un outil Multi-DB explorer



## Nos Gold Standard

- Disponibles sur Data INRAE <https://doi.org/10.15454/BVXD7I>
- 73 aliments de CIQUAL sont alignés avec des aliments SIREN

ciquali	ciqualProduct name	ciqualProduct name_en	closest foodonProductId	closest foodonProduct	closest foodon family Id	closest foodon family Inam
28300	Porc, rôti, cru	Pork, roast, raw	<a href="http://purl.obolibrary.org/obo/FOODON_03302735">http://purl.obolibrary.org/obo/FOODON_03302735</a>	pork (raw)	<a href="http://purl.obolibrary.org/obo/FOODON_00002303">http://purl.obolibrary.org/obo/FOODON_00002303</a>	swine meat food product
28002	Porc, poitrine, crue	Pork, belly, raw	<a href="http://purl.obolibrary.org/obo/FOODON_03302735">http://purl.obolibrary.org/obo/FOODON_03302735</a>	pork (raw)	<a href="http://purl.obolibrary.org/obo/FOODON_00002303">http://purl.obolibrary.org/obo/FOODON_00002303</a>	swine meat food product
28201	Porc, filet, maigre, cru	Pork tenderloin, lean, raw	<a href="http://purl.obolibrary.org/obo/FOODON_03309967">http://purl.obolibrary.org/obo/FOODON_03309967</a>	pork tenderloin (raw)	<a href="http://purl.obolibrary.org/obo/FOODON_00002303">http://purl.obolibrary.org/obo/FOODON_00002303</a>	swine meat food product
28101	Porc, côte, grillée	Pork, chop, grilled	<a href="http://purl.obolibrary.org/obo/FOODON_03311467">http://purl.obolibrary.org/obo/FOODON_03311467</a>	pork chop (cooked)	<a href="http://purl.obolibrary.org/obo/FOODON_00001049">http://purl.obolibrary.org/obo/FOODON_00001049</a>	pork chop food product
28100	Porc, côte, crue	Pork, chop, raw	<a href="http://purl.obolibrary.org/obo/FOODON_03311076">http://purl.obolibrary.org/obo/FOODON_03311076</a>	pork chop (raw)	<a href="http://purl.obolibrary.org/obo/FOODON_00002303">http://purl.obolibrary.org/obo/FOODON_00002303</a>	swine meat food product

## Nos Gold Standard

- Disponibles sur Data INRAE <https://doi.org/10.15454/BVXD71>
- 73 aliments de CIQUAL sont alignés avec des aliments SIREN
- 181 aliments de CIQUAL associés à la famille la plus proche dans SIREN

1	ciqualProduct	Nb of fa	nearestFamilyNameGoldenStandard
42	Brain, pork, braised	15	swine meat food product
43	Brain, pork, raw	15	swine meat food product
53	Chipolata slim sausage, raw	17	swine meat food product
54	Chitterling sausage from Troyes, raw	17	swine meat food product
55	Chitterling sausage, sautéed/pan-fried	18	swine meat food product
59	Cooked ham, Parisian-style, rind left on	19	swine meat food product
60	Cooked ham, superior quality, rind left on	20	swine meat food product
73	Garlic sausage	19	swine meat food product
85	Heart, pork, raw	15	swine meat food product
94	Knuckle of ham, cooked	16	swine meat food product
113	Liver, pork, raw	15	swine meat food product
135	Pork tenderloin roast, cooked	16	swine meat food product
136	Pork tenderloin, lean, raw	16	swine meat food product
137	Pork, belly, raw	15	swine meat food product
139	Pork, chop, raw	16	swine meat food product
140	Pork, roast, raw	15	swine meat food product
141	Pork, spare-ribs, raw	16	swine meat food product
153	Smoked lardoons, cooked	17	swine meat food product
154	Smoked lardoons, raw	17	swine meat food product

## Ce qui intéresse ROCED?

- Domaine d'application
  - noms d'aliments pour croiser des tables nutritionnelles
- Méthodologie de construction
  - <https://hal.archives-ouvertes.fr/hal-03372310>
- Description générale du contenu du jeu de données
  - CIQUAL <https://doi.org/10.15454/6CEYU3>
  - SIREN <https://doi.org/10.15454/5LLGVY>
  - Gold Standard <https://doi.org/10.15454/BVXD7I>
- Moyen d'accès au jeu de données: URL
- Contraintes liées à l'utilisation: CC BY 2.0
- Caractérisation du jeu de données du point de vue raisonnement (temporel, spatial, incertitude, ...)
  - Famille la plus proche