

Matching Food Items with FoodOn Identifiers in a Knowledge Base

Kat Thornton, Kenneth Seals-Nutt, Mika Matsuzaki

15 August 2022

Our Collaboration



Figure 1: Mika Matsuzaki- nutritional epidemiologist,
Kenneth Seals-Nutt- software engineer, **Kat Thornton-** information scientist

Will the combination of data sourced from global food composition tables and enriched with FoodOn identifiers reduce data disparities for people who record nutrient intake?

Wikipedias



WIKIPEDIA
The Free Encyclopedia

Wikimedia Foundation Projects

Wikipedia

Wiktionary

Wikiquote

Wikibooks

Wikisource

Wikispecies

Wikinews

Wikiversity

Wikivoyage

Commons

Wikidata

MediaWiki

Meta-Wiki

Incubator

Labs

Foundation



WIKIDATA

Wikibase

Wikibase is the software used by Wikidata that lets us store, manage, and access structured data using peer production.



Goal of the project

To create a free, open, dynamic, and collaborative global food composition database using a peer production model.

We have been creating our data models using the following examples.

[Visit WikiFCD!](#)

- USDA Food Data Central
- SMILING Vietnam
- SMILING Indonesia
- SMILING Laos
- SMILING Thailand
- FCT Malawi
- FCT Nigeria
- FCT Uganda

Food Composition Table Examples

Table 1. Proximate composition, content of cholesterol and phytonutrients and the proportion of fatty acids of food samples from Ghana

Food table no.	Food	Lab code	Proximate composition, weight/100 g edible portion						Cholesterol mg	Phytonutrol mg
			Water g	Protein g	Fat g	Total carbohydrate g				
Cereal products										
10	akassa, koko	554	93.8	0.5	0.4	5.3	—	—	—	
11	akpler	586	66.6	1.4	0.8	31.2	—	—	—	
13	banku	555	70.6	2.7	1.3	25.4	—	—	—	
19	kenkey, Ga	556	64.2	3.6	1.7	30.5	—	—	—	
20	kenkey, Fante	557	70.5	3.0	1.4	25.1	—	—	—	
29	rice, raw	558	8.4	8.1	1.0	82.5	—	—	—	
31	rice, cooked	559	64.7	2.9	0.4	32.0	—	—	—	
35	wheat bread	560	24.3	11.7	3.7	60.3	—	—	—	
37	toughi	573	28.0	9.5	6.9	55.6	—	—	—	
350	biscuits	582	35.4	8.8	6.7	49.1	—	—	—	
429	ricewater	609	90.6	1.0	0.3	8.1	—	—	—	
434	wheat	610	62.4	4.6	1.8	31.2	—	—	—	
Starchy roots and fruits										
39	cassava, raw	561	57.2	0.7	0.5	41.6	—	—	—	
40	cassava, boiled	587	59.6	0.7	0.5	39.2	—	—	—	
44	cassava balls, fried	588	19.5	0.7	6.8	73.0	—	—	—	
46	garri	562	7.7	0.9	1.1	90.3	—	—	—	
47	kokonte	563	71.1	0.5	0.5	27.9	—	—	—	
49	yake yake	589	53.2	0.4	0.5	45.9	—	—	—	
53	cocoyam, boiled	564	59.2	3.0	0.6	37.2	—	—	—	

U.S. DEPARTMENT OF AGRICULTURE
Agricultural Research Service

[HOME](#)
[DATA TYPE DOCUMENTATION](#)
[DOWNLOAD DATA](#)
[API GUIDE](#)
[HELP](#)
[FAQ](#)
[ABOUT US](#)
[CONTACT FOODDATA CENTRAL](#)

FoodData Central

Search FoodData Central

[Download Data](#)

[Get an API Key](#)

[API Guide](#)

```

$ curl -X GET 'https://api.nal.usda.gov/fdc/v1/food/1103278?api_key=DEMO_KEY' -H 'accept: application/json' | jq
  % Total    % Received % Xferd  Average Speed   Time    Time     Current
                                 Dload  Upload   Total   Spent    Left  Speed
100 12973  0 12973  0    0 6878   0 --:--:--  0:00:01 --:--:-- 6878

{
  "value": {
    "id": 977039,
    "value": "from restaurant",
    "sequenceNumber": 1,
    "foodAttributeType": {
      "id": 1001,
      "name": "Additional Description",
      "description": "Additional descriptions for the food."
    }
  },
  "value": "with or without fat",
  "sequenceNumber": 1,
  "foodAttributeType": {
    "id": 1001,
    "name": "Additional Description",
    "description": "Additional descriptions for the food."
  }
}

```

FOOD_NAME_ENGLISH	FOOD_NAME_LOCAL	SCIENTIFIC_NAME	FOOD_GROUP	FOOD_SUB_GROUP	COUNTRY	LA
1	Rice noodle, fermented	Khao jeon	Grains & grain products	Refined grains and products, unenriched/unrefined	Laos	01
2	White noodle (baked), instant	Khao jeon-let baob	Grains & grain products	Refined grains and products, unenriched/unrefined	Laos	01
3	Sticky rice (white), steamed	Khao niao reung	Grains & grain products	Refined grains and products, unenriched/unrefined	Laos	01
4	Rice steamed, white	Khao khao reung	Grains & grain products	Refined grains and products, unenriched/unrefined	Laos	01
5	Rice noodle (mass sold)	Son-Khao pick	Grains & grain products	Refined grains and products, unenriched/unrefined	Laos	01
6	Rice, white, steamed	Khao khao	Grains & grain products	Refined grains and products, unenriched/unrefined	Laos	01
7	Rice, white, fried	Khao khao khao	Grains & grain products	Refined grains and products, unenriched/unrefined	Laos	01
8	Rice, white	Khao khao	Grains & grain products	Refined grains and products, unenriched/unrefined	Laos	01
9	Sticky rice (white), steamed	Khao khao reung	Grains & grain products	Refined grains and products, unenriched/unrefined	Laos	01
10	Baby food, vegetable	Maach daek ka Geran	Grains & grain products	Refined grains and products, unenriched/unrefined	Laos	01
11	Peas, cooked	Thoum ka loun	Legumes, nuts & seeds	Cooked beans, lentils, peas	Laos	02
12	Cornmeal, white, dried, instant	Thoum ka loun	Legumes, nuts & seeds	Cooked beans, lentils, peas	Laos	02
13	Bean, white, fresh	Maach pae	Legumes, nuts & seeds	Cooked beans, lentils, peas	Laos	02
14	Bean, white, fresh	Maach pae	Legumes, nuts & seeds	Cooked beans, lentils, peas	Laos	02
15	Bean, white, fresh	Maach pae	Legumes, nuts & seeds	Cooked beans, lentils, peas	Laos	02
16	Bean, white, fresh	Maach pae	Legumes, nuts & seeds	Cooked beans, lentils, peas	Laos	02
17	Bean, white, fresh	Maach pae	Legumes, nuts & seeds	Cooked beans, lentils, peas	Laos	02
18	Bean, white, fresh	Maach pae	Legumes, nuts & seeds	Cooked beans, lentils, peas	Laos	02
19	Bean, white, fresh	Maach pae	Legumes, nuts & seeds	Cooked beans, lentils, peas	Laos	02
20	Bean, white, fresh	Maach pae	Legumes, nuts & seeds	Cooked beans, lentils, peas	Laos	02

Example FCT

Table 4. Carotenoids

Food code	Food Name	No. of Regions	Lutein	Zeaxanthin	Lycopene	β -Cryptoxanthin	γ -Carotene	α -Carotene	β -Carotene	Total Carotenoids
			μg							
			LUTN	ZEAX	LYCPN	CRYPXB	CARTG	CARTA	CARTB	CARTD
E020	Fig (<i>Ficus carica</i>)	6	19.0 \pm 2.27	2.01 \pm 0.64					2.40 \pm 0.59	454 \pm 51.3
E021	Gooseberry (<i>Embelica officinalis</i>)	5	38.7 \pm 9.78	2.86 \pm 1.61					1.58 \pm 0.09	62.01 \pm 6.45
E022	Grapes, seeded, round, black (<i>Vitis vinifera</i>)	4	27.8 \pm 8.85	3.59 \pm 2.20					29.36 \pm 8.37	213 \pm 29.4
E023	Grapes, seeded, round, green (<i>Vitis vinifera</i>)	5	31.3 \pm 8.77	4.06 \pm 1.60					30.77 \pm 11.10	208 \pm 28.8
E024	Grapes, seeded, round, red (<i>Vitis vinifera</i>)	5	17.2 \pm 3.04	7.58 \pm 2.28					19.94 \pm 1.29	314 \pm 99.9
E025	Grapes, seedless, oval, black (<i>Vitis vinifera</i>)	5	19.4 \pm 3.07	8.14 \pm 1.73					19.73 \pm 1.76	321 \pm 50.6
E026	Grapes, seedless, round, green (<i>Vitis vinifera</i>)	5	25.8 \pm 7.90	6.82 \pm 3.77					25.46 \pm 6.66	216 \pm 31.3
E027	Grapes, seedless, round, black (<i>Vitis vinifera</i>)	5	16.1 \pm 2.90	7.95 \pm 1.73					20.58 \pm 2.36	232 \pm 27.9
E028	Guava, white flesh (<i>Psidium guajava</i>)	5	19.7 \pm 4.59	5.92 \pm 3.06					298 \pm 32.8	996 \pm 108
E029	Guava, pink flesh (<i>Psidium guajava</i>)	5	90.3 \pm 11.5	8.94 \pm 2.21	2843 \pm 340				267 \pm 88.2	4078 \pm 477
E030	Jack fruit, ripe (<i>Artocarpus heterophyllus</i>)	5	19.3 \pm 6.91	2.42 \pm 1.27					23.53 \pm 1.24	59.61 \pm 4.90
E031	Jambu fruit, ripe (<i>Syzygium cumini</i>)	2	18.49	5.68					1.55	66.50
E032	Karonda fruit (<i>Carissa carandas</i>)	1	6.12	1.14					15.64	55.89
E033	Lemon, juice (<i>Citrus limon</i>)	6	9.68 \pm 2.22	1.72 \pm 0.50					2.62 \pm 0.34	85.99 \pm 3.36
E034	Lime, sweet pulp (<i>Citrus limetta</i>)	6	13.2 \pm 3.54	1.89 \pm 0.44					2.54 \pm 0.30	85.13 \pm 10.71
E035	Litchi (<i>Nephelium litchi</i>)	4	27.33 \pm 9.81	1.61 \pm 0.62					1.47 \pm 0.21	129 \pm 15.1
E036	Mango, ripe, banganapalli (<i>Mangifera indica</i>)	6	3.19 \pm 1.51	1.62 \pm 0.22		2.72 \pm 1.31			1168 \pm 141	1424 \pm 308
E037	Mango, ripe, gulabkhas (<i>Mangifera indica</i>)	2	4.18	1.64		4.58			666	1350
E038	Mango, ripe, himsagar (<i>Mangifera indica</i>)	3	6.26 \pm 2.09	1.19 \pm 0.14		5.60 \pm 3.84			1181 \pm 136	1423 \pm 201
E039	Mango, ripe, kesar (<i>Mangifera indica</i>)	4	3.71 \pm 2.21	2.33 \pm 0.17		6.96 \pm 3.31			1264 \pm 172	1438 \pm 314
E040	Mango, ripe, neelam (<i>Mangifera indica</i>)	2	2.95	1.67		2.51			1291	1376

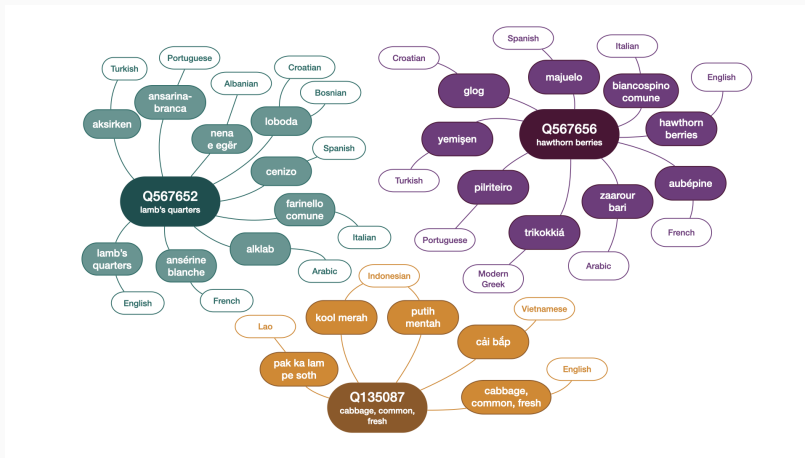
102

Matching Food Items to FoodOn Identifiers


- For each new FCT we add we need to map hundreds of food items to FoodOn identifiers

Sample_Id	Sample_Desc	Processed_Sample	Processed_Sample (With Scientific Name)	Matched_Components	Match_Status(Macro Level)
1	Acha, black, whole grain, Raw	acha black whole grain raw	acha black whole grain raw	['food (raw):FOODON_03311126', 'whole grain:FOODON_00003950']	Component Match
2	Acha, white, whole grain, Raw	acha white whole grain raw	acha white whole grain raw	['food (raw):FOODON_03311126', 'white:PATO_0000323', 'whole grain:FOODON_00003950']	Component Match
3	Acha, white, whole grain, Boiled	acha white whole grain boiled	acha white whole grain boiled	['food (boiled):FOODON_00002688', 'white:PATO_0000323', 'whole grain:FOODON_00003950']	Component Match
4	Maize, white, whole kernel, dried, raw	maize white whole kernel dried raw	maize {zea mays} white whole kernel dried raw	['food (dried):FOODON_03307539', 'food (raw):FOODON_03311126', 'maize kernel:FOODON_00003427', 'seed, skin present, germ present:FOODON_03420133', 'white:PATO_0000323']	Component Match

Multilingual Data in WikiFCD



How FoodOn Helps WikiFCD



- Main page
- Recent changes
- Random page
- Help about MediaWiki

Tools

- Special pages
- Printable version

Wikibase

- New Item
- New Property
- New Schema
- All Properties
- Query Service
- Cradle
- QuickStatements

English Create account Log in

Special page Search wikifcd

Search results

Advanced search: Sort by relevance X

Search in: (Main) X (Item) X

Tomato, fresh (Q135084)
edible fruit
57 statements, 0 sitelinks - 17:12, 7 January 2021

MICROWAVEABLE TOMATO BASIL SOUP, TOMATO BASIL (Q230096)
Food Item with USDA ID 735160
16 statements, 0 sitelinks - 23:21, 12 November 2020

TOMATO JUICE FROM CONCENTRATE, TOMATO (Q516975)
Food Item with USDA ID 707404
17 statements, 0 sitelinks - 03:24, 26 November 2020

Results 1 - 21 of 3,929

Demo



Thank you!

katherine.thornton@yale.edu