Knowledge Engineering in Food Computing – Selected Problems and Applications

Weronika T. Adrian, Ph.D.

KRaKEn' Research Group | AGH University of Kraków

International Joint Conference on Rough Sets

October 7th 2023

Overview of the talk

- Background, context and motivations
  - Knowledge engineering
  - Food computing
  - TAISTI project
- Ontological modeling in the food domain
- Knowledge-based reasoning about substitutes
- Open challenges and opportunities

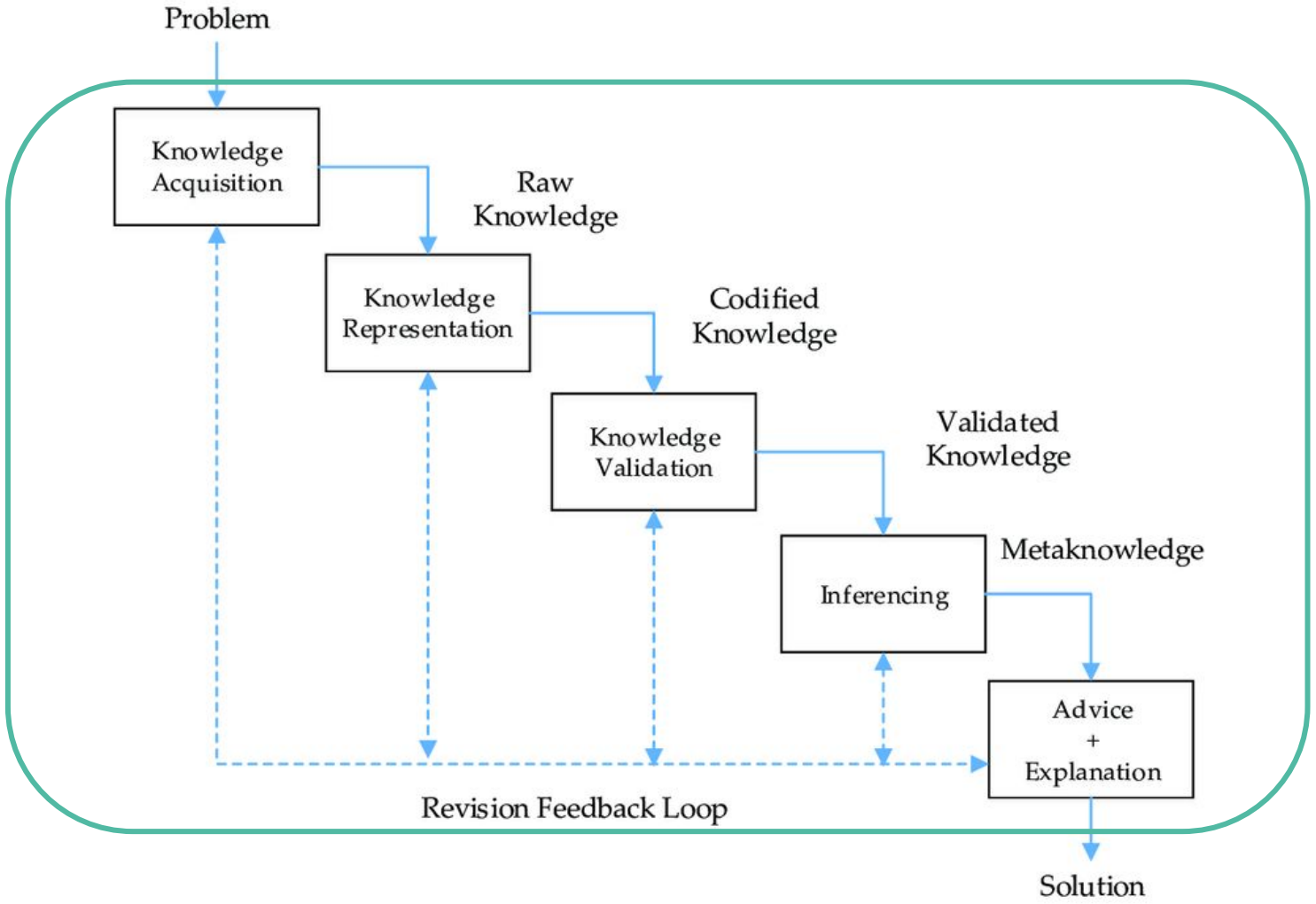
…The best ideas shape at the crossroads of domains…

Introduction

What is knowledge?

Go to: Menti.com
and enter code:
4236 7543

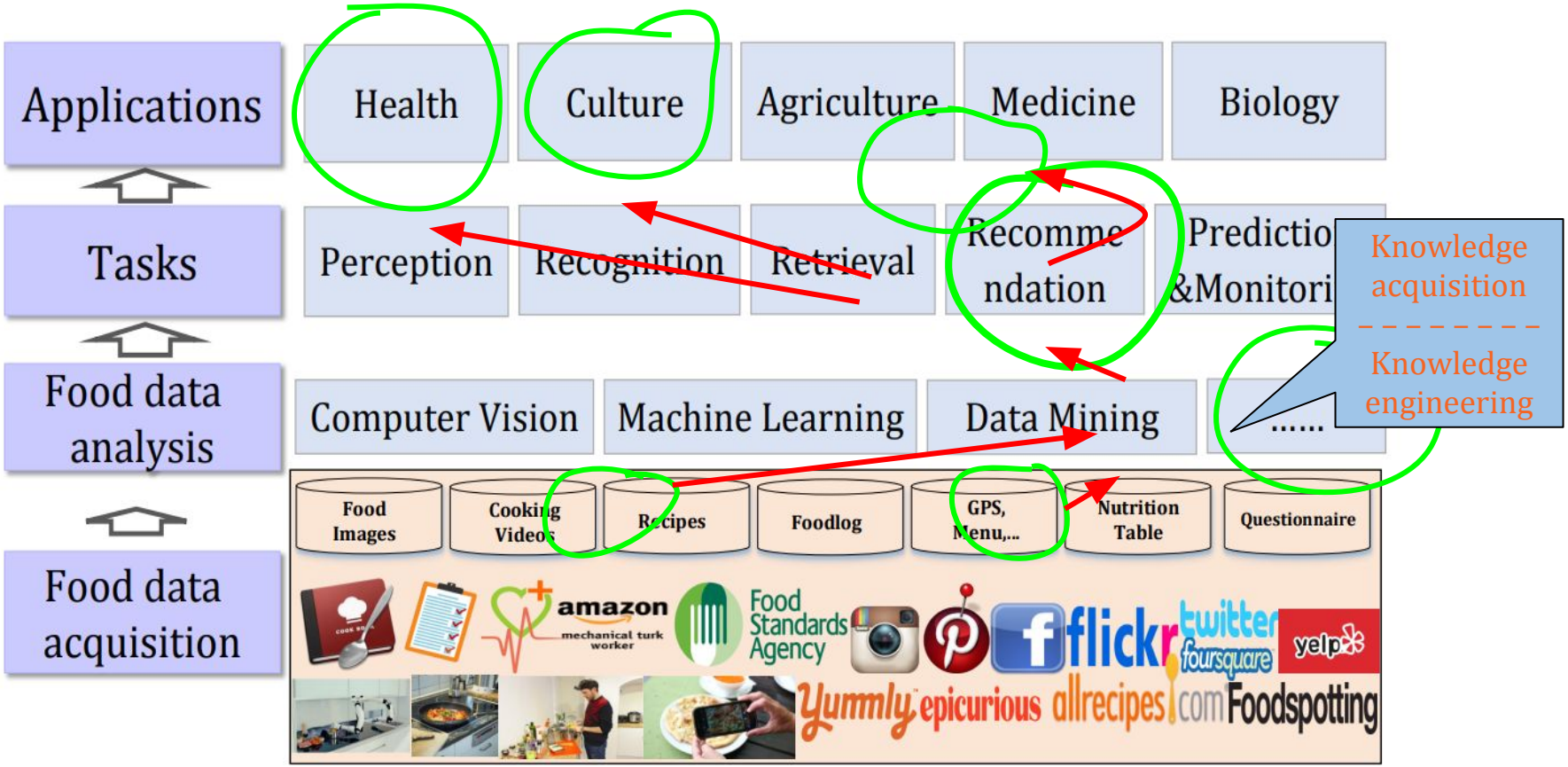
What is knowledge engineering?


Problem
Knowledge Acquisition
Raw Knowledge
Knowledge Representation
Codified Knowledge
Knowledge Validation
Validated Knowledge
Inferencing
Metaknowledge
Advice + Explanation
Revision Feedback Loop
Solution

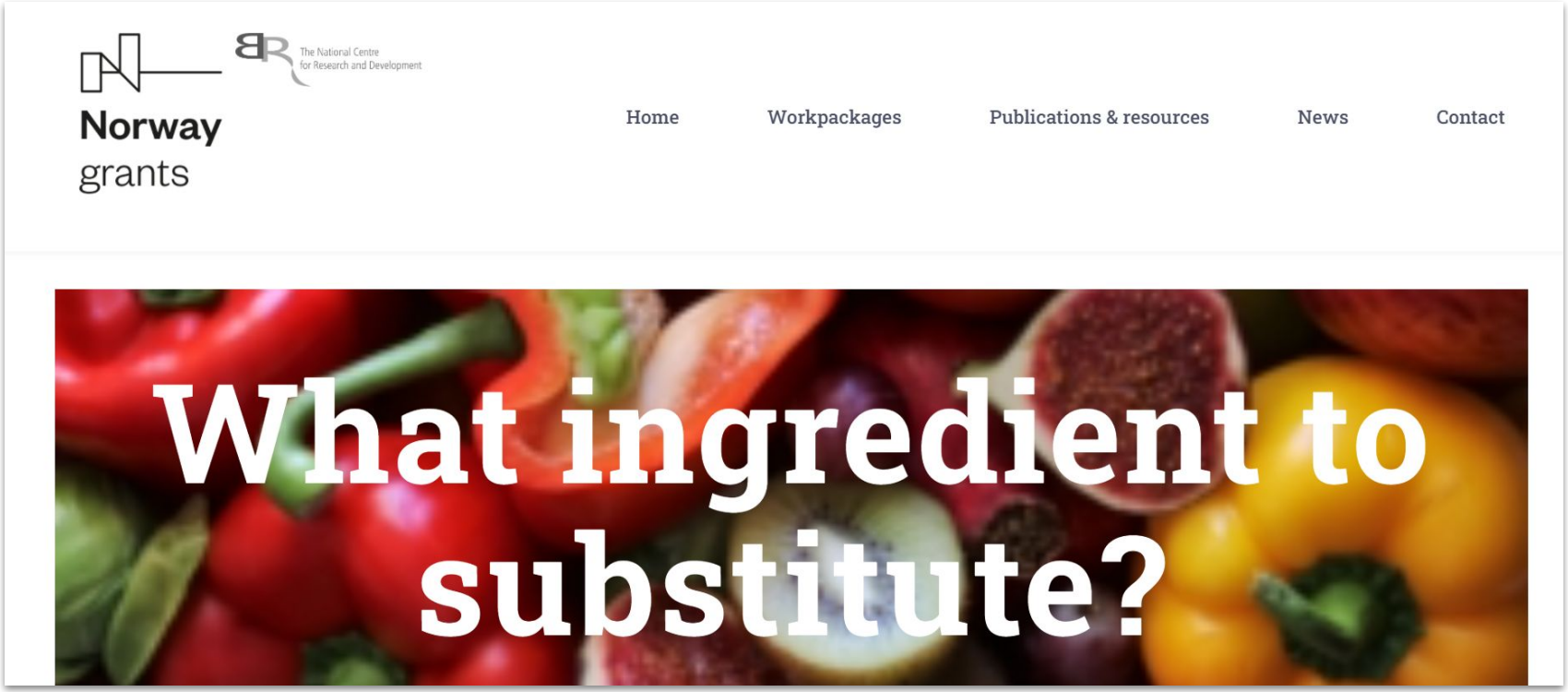
What is food?

Again:
menti.com
4236 7543

What is food computing?



TAISTI: AI-based substitution recommendation



TAISTI in a nutshell (http://taisti.eu)

- “Development of a **T**echnology based on **A**rtificial **I**ntelligence for inferring **S**ubs**T**itutable recipe **I**ngredients”

- EEA and Norway grants - Financial contribution of Iceland, Liechtenstein and Norway
- Programme „Applied Research”
- Managed by the National Centre for Research and Development
- Small grant scheme: 2 year project
- 856 994 PLN in funding

Multidisciplinary research team

- Leader, PI: **Agnieszka Ławrynowicz** (Computer Science, Poznań)
  - Jędrzej Potoniec (Computer Science - **Ontologies, KKR**, Poznań)
  - Dawid Wiśniewski (Computer Science - **IE, entity linking**, Poznań)
- Agnieszka Kalicka (**Linguistics**, Poznań)
- Anna Gramza-Michałowska (**Food technology**, Poznań)
- Bartosz Kulczyński (**Dietetics**, Poznań)
- Anna Wróblewska (Computer Science - **recommenders, ML**, Warszawa)
- Weronika T. Adrian (Computer Science - **logic-based AI, KE**, Kraków)
- Phd Students, Master Students, young researchers

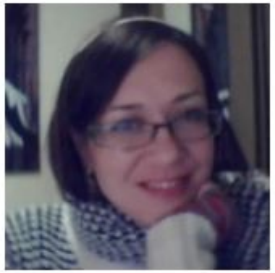
Multidisciplinary research team

Agnieszka Ławrynowicz

Anna Wróblewska

Weronika T. Adrian

Dawid Wiśniewski



Agnieszka Kaliska

Anna Gramza-Michałowska

and
Jędrzej Potoniec, Mateusz Lango, Witold Sosnowski, Maciej Pawłowski, Bartosz Kulczyński, Sergiy Tkachuk, Adam Lewandowski, Andrzej Gretkowski, Maciej Kutyła, Jakub Dutkiewicz

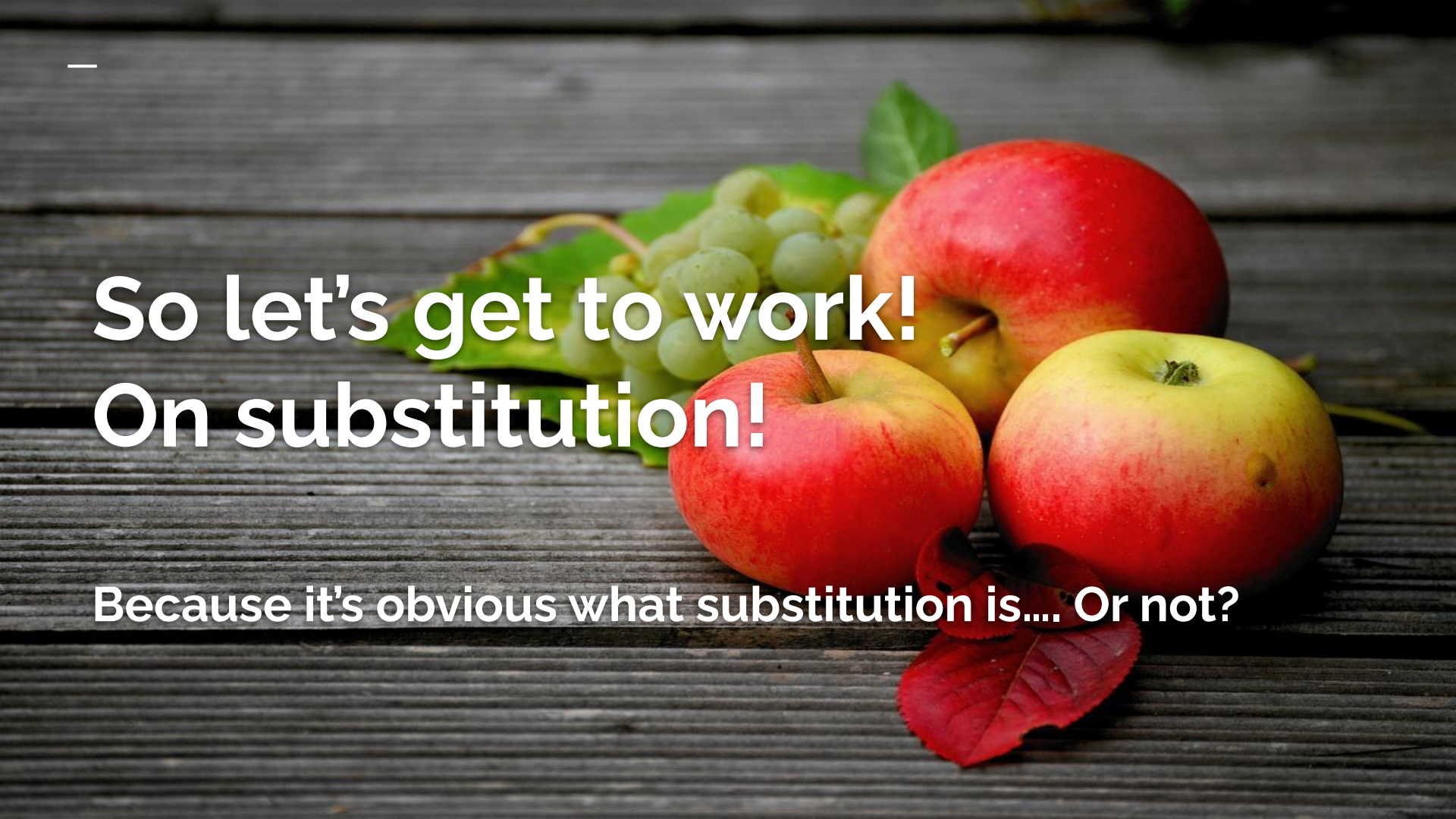
R&D problems and challenges

1. **Knowledge curation**
   a. Manual annotation (linguistics)
   b. Entity linking (information extraction, linking instances to existing categories and objects)
2. **Knowledge acquisition, integration and modeling**
   a. Existing sources: recipes, nutritional tables, ontologies
   b. Domain experts (translating unstructured informal knowledge into knowledge structured and formal)
3. *Machine learning-based recommendations*
4. **Rule-based verification and explanation**

Knowledge engineering questions

1. Where/What is the useful knowledge?
2. How to conceptualize/integrate/model it?
3. How to reason about it?

Knowledge analysis and acquisition


So let's get to work!
On substitution!
Because it's obvious what substitution is.... Or not?

What is substitution?

Answers food technologist…

Substitute: ingredient(s) change, same technological process, final product similar to the original one

Analogue (imitation): ingredient(s) change, technological process may differ, the final product imitates the original one in terms of color, taste etc.

Domain experts' sources…

Gluten w spożywczych zastosowaniach pozapiekarskich

Typ produktu	Zawartość glutenu %	Typ produktu	Zawartość glutenu %
Analogi owoców morza	1.3	Restrukturyzowane steki wołowe	3.6
Analogi mięsa krabiego	2.1	Frankfurterki	8
Analogi kawioru	1-30	Przekąski wysokobiałkowe	1-50
Analogi kiełbasek	8-17	Pasty wysokobiałkowe	1.6
Analogi pulpetów i hamburgerów	10.6	Tortille	1-4
Imitacje serów	5.8-14.2	Ekstrudowane produkty błonnikowe z glutenem	20-23

HOFFMANN M, JĘDRZEJCZYK H. NOWE ANALOGI MIĘSA. POSTĘPY TECHNIKI PRZETWÓRSTWA SPOŻYWCZEGO 1/2010

		(56): Podano w przeliczeniu na glin.			

▼M2

12.1.2 Substytuty soli

Grupa I	Dodatki			
E 338–452	Kwas fosforowy – fosforany – di-, tri- i polifosforany	10 000	(1) (4)	

▼M2

Numer kategorii	Numer E	Nazwa	Maksymalny poziom (odpowiednio mg/l lub mg/kg)	Przypisy	Ograniczenia/wyjątki
	E 535–538	Żelazocyjanki	20	(1) (57)	
▼M7	E 551–559	Dwutlenek krzemu – krzemiany	20 000		Okres stosowania: do dnia 31 stycznia 2014 r.
	E 551–553	Dwutlenek krzemu – krzemiany	20 000		Okres stosowania: od dnia 1 lutego 2014 r.
▼M2	E 620–625	Kwas glutaminowy – glutaminiany	*quantum satis*		
	E 626–635	Rybonukleotydy	*quantum satis*		
		(1): Dodatki mogą być dodawane pojedynczo lub łącznie.			
		(4): Maksymalny poziom podano w przeliczeniu na P_2O_5.			
		(57): Maksymalny poziom podano w przeliczeniu na bezwodny żelazocyjanek potasu.			

[**Scenario #1** - zamiana wynikająca z braku produktu; zamiennik musi odwzorowywać wartość odżywczą, tj. być również dobrym źródłem błonnika]

Persona X zdecydowała się na przygotowanie risotto. Jednak, gdy przyszło do sporządzenia potrawy okazało się, że zabrakło jej ryżu brązowego. W przepisie, który wykorzystywała nie było podanej informacji na temat składnika/składników, którymi można zastąpić ryż. Co więcej Persona X postanowiła odżywiać się zdrowo i zastanawia się, który możliwy zamiennik ryżu spełniałby jej oczekiwania, tj. nadawałby się nie tylko z technologicznego punktu widzenia, ale przede wszystkim byłby dobrym źródłem błonnika pokarmowego.

[**Scenario #2** - zamiana wynikająca z konieczności wykluczenia konkretnego produktu z powodu zdrowotnego; zamiennik musi spełniać wymagania zdrowotne]

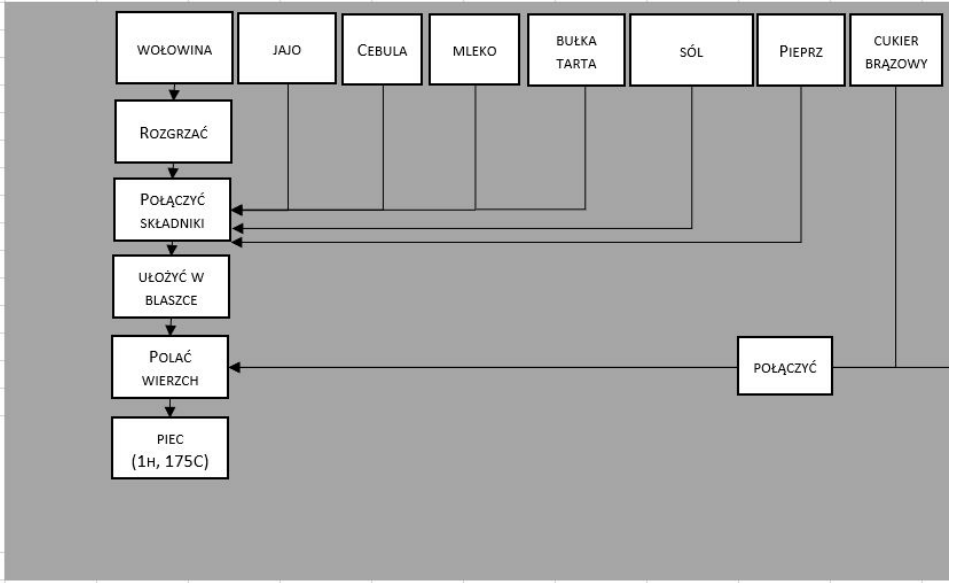
Persona Y postanowiła przyrządzić na imprezę urodzinową deser mleczny z owocami. Jednak okazało się, że jeden z gości ma zdiagnozowaną alergię na białko mleka krowiego i orzechy. Aby całkowicie nie rezygnować z pomysłu przygotowania słodkiej przekąski, Persona Y postanowiła wymienić mleko na inny składnik. Niestety nie ma ona pomysłu, który produkt sprawdziłby się w tym przypadku jako dobry zamiennik mleka. Mąż podsunął jej pomysł, że może to być napój migdałowy. Jednak nie jest przekonana, czy będzie on odpowiednim zamiennikiem, gdyż wspomniany gość ma zdiagnozowaną alergię na orzechy.

[**Scenario #3** - zamiana wynikająca z konieczności wykluczenia konkretnego produktu z powodu zdrowotnego; zamiennik musi spełniać wymagania technologiczne, tj. musi nadawać taką samą słodkość potrawie, jak cukier]

Persona Z jest w trakcie przygotowania pieczonych ciasteczek na spotkanie rodzinne. Dowiedziała się jednak

Domain experts' analyses…

B	C
https://www.allrecipes.com/recipe/16354/easy-meatloaf/	
Easy Meatloaf / mielony - pieczeń rzymska	
1 ½ pounds ground beef	mielona wołowina
1 egg	jajo
1 onion, chopped	cebula
1 cup milk	mleko
1 cup dried bread crumbs	bułka tarta / krakersy
salt and pepper to taste	sól i pieprz
2 tablespoons brown sugar	brązowy cukier
2 tablespoons prepared mustard	musztarda
⅓ cup ketchup	keczup
Step 1	
Preheat oven to 350 degrees F (175 degrees C).	
Step 2	
In a large bowl, combine the beef, egg, onion, milk and bread OR cracker crumbs. Season with salt and pepper to taste and place in a lightly greased 9x5-inch loaf pan, or form into a loaf and place in a lightly greased 9x13-inch baking dish.	
Step 3	
In a separate small bowl, combine the brown sugar, mustard	



Pieczeń rzymska mielony SCHEMAT | Pieczeń - ZAMIENNIKi | Wieprzowina- KARTA PRODUKTU | ...

Domain experts' analyses...

SKŁADNIKI PODSTAWOWE	ZAMIEŃ NA:	WARUNEK 1:	WARUNEK 2:	WARUNEK 3 [STOSUNEK ZAMIANY]:	
wołowina mielona	wieprzowina	na zimno: TAK	pieczenie: TAK	1:1	
	kurczak	na zimno: TAK	pieczenie: TAK	1:1	
	indyk	na zimno: TAK	pieczenie: TAK	1:1	
	łosoś biały	na zimno: TAK	pieczenie: TAK	1:1	
	tuńczyk	na zimno: TAK	pieczenie: TAK	1:1	
	ciecierzyca	ugotowana	pieczenie: TAK	1:1	
	soczewica	ugotowana	pieczenie: TAK	1:1	
	tempeh	na zimno: TAK	pieczenie: TAK	1:1	
	soja	ugotowana	pieczenie: TAK	1:1	
	sojowe kotlety	na zimno: TAK	pieczenie: TAK	1:1	
	fasola	ugotowana	pieczenie: TAK	1:1	
	grzyby	na zimno: TAK	pieczenie: TAK	1:1	
	Seitan (gluten pszenny)	na zimno: TAK	pieczenie: TAK	1:1	
	białko z owadów	na zimno: TAK	pieczenie: TAK	1:1	
	bakłażan	na zimno: TAK	pieczenie: TAK	1:1	
jajo	jajo w proszku	na zimno: TAK	pieczenie: TAK	1:1	
	len mielony	na zimno: TAK	pieczenie: TAK	1:1	plus woda 100%
	chia mielona	na zimno: TAK	pieczenie: TAK	10g:1szt	plus woda 100%
	agar agar	na zimno: TAK	pieczenie: TAK	5g:1 szt	plus woda 100%

Pieczeń rzymska mielony SCHEMAT | **Pieczeń - ZAMIENNIKi** | Wieprzowina- KARTA PRODUKTU | ...

Domain experts' analyses…

Wieprzowina	zależy który element kulinarny

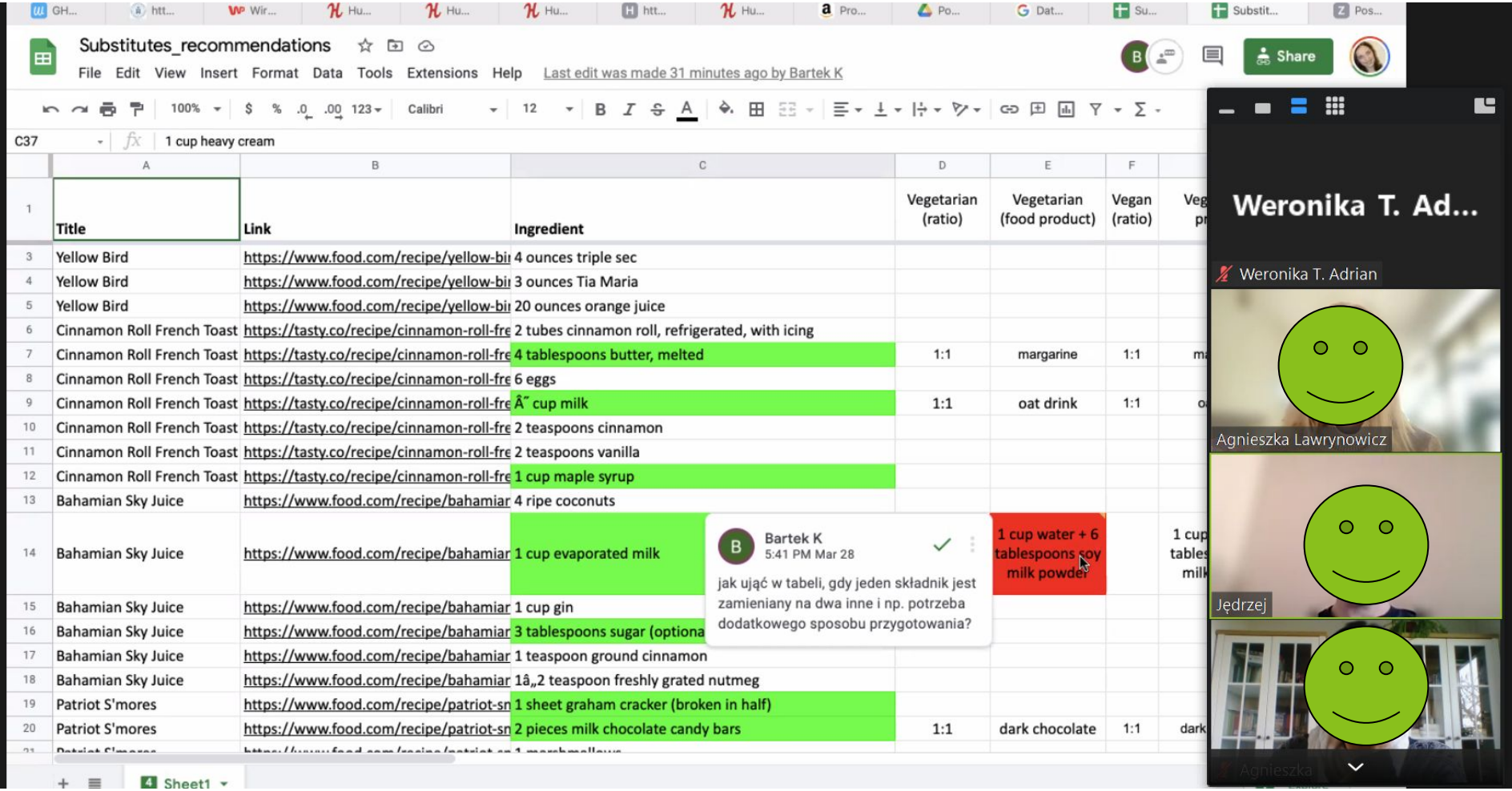
ETYKIETY OGÓLNE:	LEKKOSTRAWNY	BEZGLUTENOWY	WEGETARIAŃSKI	WEGAŃSKI	BEZ LAKTOZY
	NIE	TAK	NIE	NIE	TAK

ETYKIETY TECHNOLOGICZNE:	NA SUROWO	DO GOTOWANIA	DO SMAŻENIA	DO PIECZENIA
	NIE	TAK	TAK	TAK

ETYKIETY ODŻYWCZE:	ŹRÓDŁO: żelazo, białko	MAŁO: witaminy	KALORIE: 120

MOŻLIWA ZAMIANA NA:	wieprzowina	kurczak	indyk	łosoś biały	tuńczyk	ciecierzyca	soczewica	tempeh
warunek 1 [stosunek zamiany]:	1:1	1:1	1:1	1:1	1:1	1:1	1:1	
warunek 2:	na zimno: TAK	na zimno: TAK	na zimno: TAK	na zimno: TAK	na zimno: TAK	ugotowana	ugotowana	na zimn
warunek 3:								
warunek 4:								
warunek 5:								
UWAGA DODATKOWA/KOMENTARZ:								

Pieczeń rzymska mielony SCHEMAT | Pieczeń - ZAMIENNIKi | **Wieprzowina- KARTA PRODUKTU** | …


Substitutes_recommendations
File Edit View Insert Format Data Tools Extensions Help
Last edit was made 31 minutes ago by Bartek K
Share
C37
1 cup heavy cream
Title
Link
Ingredient
Vegetarian (ratio)
Vegetarian (food product)
Vegan (ratio)
Yellow Bird https://www.food.com/recipe/yellow-bi 4 ounces triple sec
Yellow Bird https://www.food.com/recipe/yellow-bi 3 ounces Tia Maria
Yellow Bird https://www.food.com/recipe/yellow-bi 20 ounces orange juice
Cinnamon Roll French Toast https://tasty.co/recipe/cinnamon-roll-fre 2 tubes cinnamon roll, refrigerated, with icing
Cinnamon Roll French Toast https://tasty.co/recipe/cinnamon-roll-fre 4 tablespoons butter, melted 1:1 margarine 1:1
Cinnamon Roll French Toast https://tasty.co/recipe/cinnamon-roll-fre 6 eggs
Cinnamon Roll French Toast https://tasty.co/recipe/cinnamon-roll-fre Â˜ cup milk 1:1 oat drink 1:1
Cinnamon Roll French Toast https://tasty.co/recipe/cinnamon-roll-fre 2 teaspoons cinnamon
Cinnamon Roll French Toast https://tasty.co/recipe/cinnamon-roll-fre 2 teaspoons vanilla
Cinnamon Roll French Toast https://tasty.co/recipe/cinnamon-roll-fre 1 cup maple syrup
Bahamian Sky Juice https://www.food.com/recipe/bahamiar 4 ripe coconuts
Bahamian Sky Juice https://www.food.com/recipe/bahamiar 1 cup evaporated milk
Bartek K
5:41 PM Mar 28
jak ująć w tabeli, gdy jeden składnik jest zamieniany na dwa inne i np. potrzeba dodatkowego sposobu przygotowania?
1 cup water + 6 tablespoons soy milk powder
Bahamian Sky Juice https://www.food.com/recipe/bahamiar 1 cup gin
Bahamian Sky Juice https://www.food.com/recipe/bahamiar 3 tablespoons sugar (optiona
Bahamian Sky Juice https://www.food.com/recipe/bahamiar 1 teaspoon ground cinnamon
Bahamian Sky Juice https://www.food.com/recipe/bahamiar 1â„2 teaspoon freshly grated nutmeg
Patriot S'mores https://www.food.com/recipe/patriot-sn 1 sheet graham cracker (broken in half)
Patriot S'mores https://www.food.com/recipe/patriot-sn 2 pieces milk chocolate candy bars 1:1 dark chocolate 1:1
Sheet1
Weronika T. Ad...
Weronika T. Adrian
Agnieszka Lawrynowicz
Jędrzej

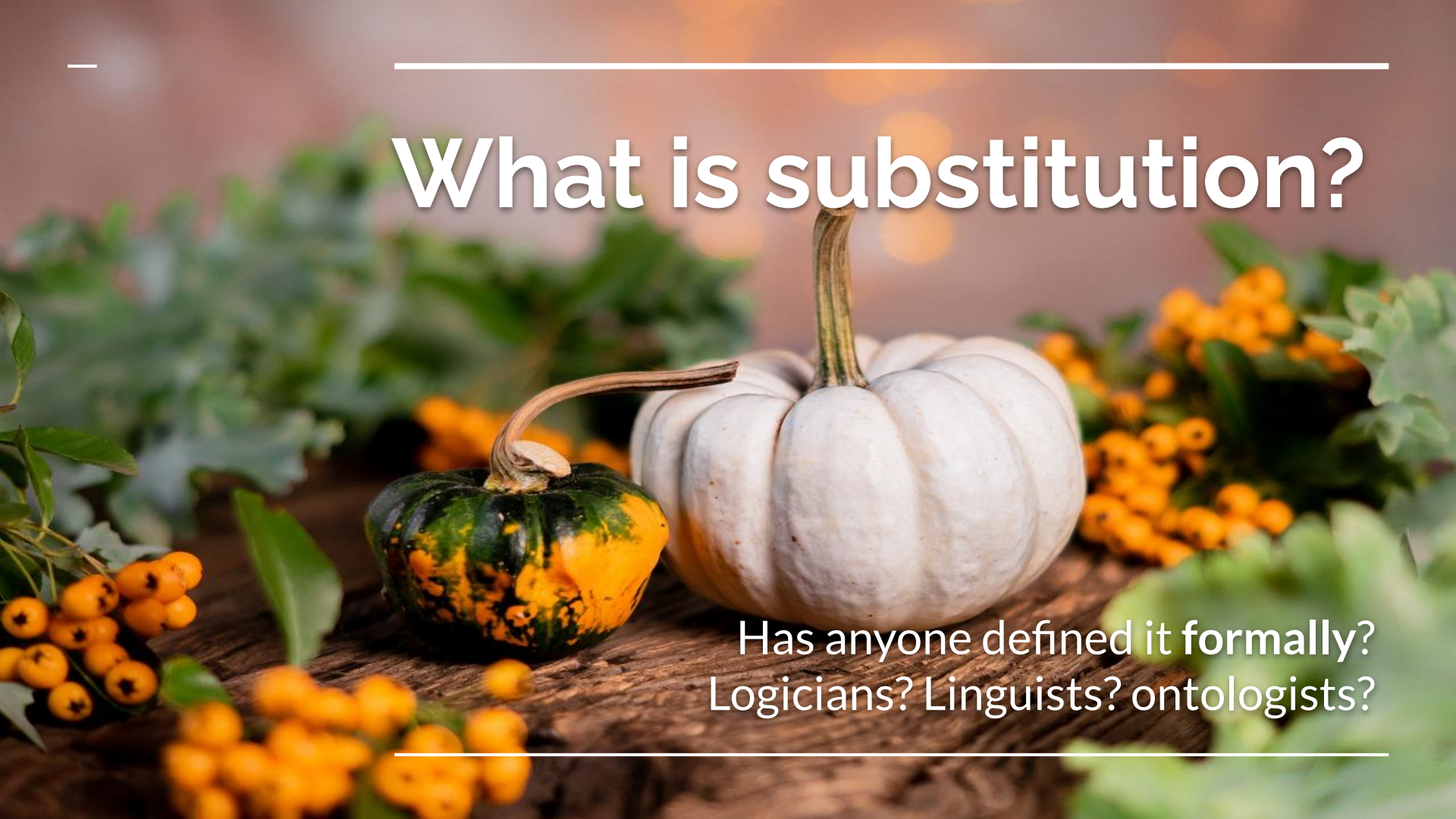
What is substitution?

Answers an ML engineer…

Something that we learned that appears in similar context as the original ingredient

But is similarity enough?

- People search for substitutes for various reasons: *allergies*, *diets*, *shortage of ingredients*
- Machine learning learns about *similar recipes*
- **But if we want to substitute something, it is because the original recipes was “wrong” in some respect, so the new recipe should be somehow similar, also different!**
- So how to model the substitution and reason about it?



What is substitution?

Has anyone defined it **formally**?
Logicians? Linguists? ontologists?

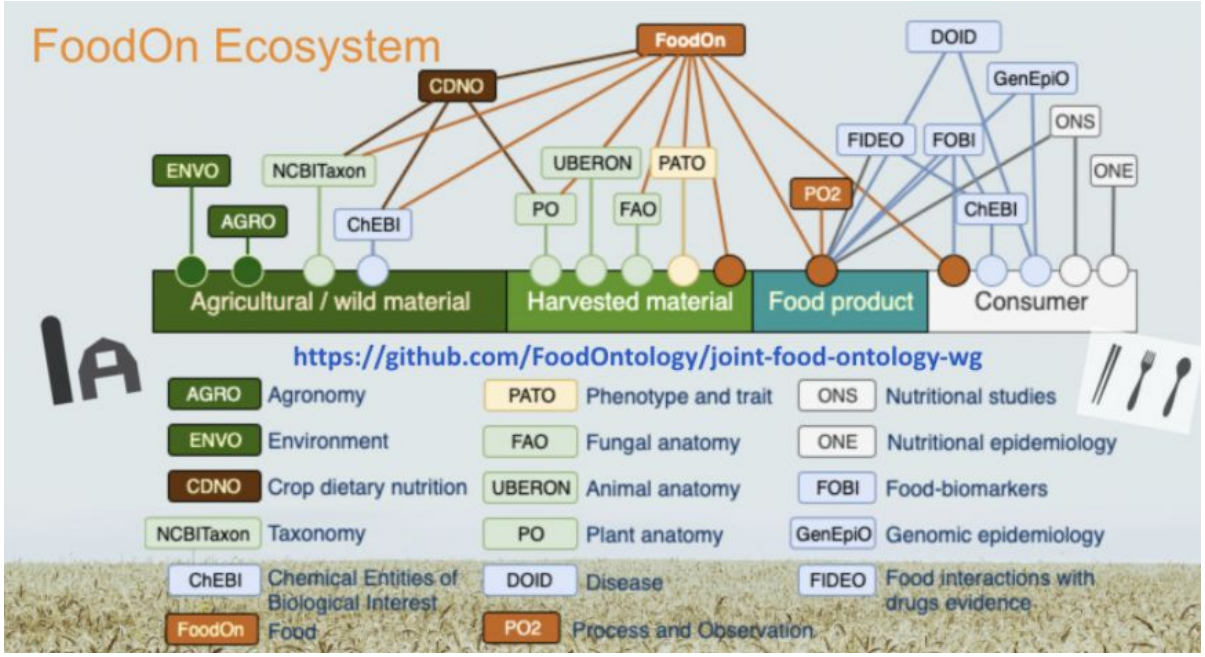
Review - substitutes

File Edit View Insert Format Data Tools Extensions Help

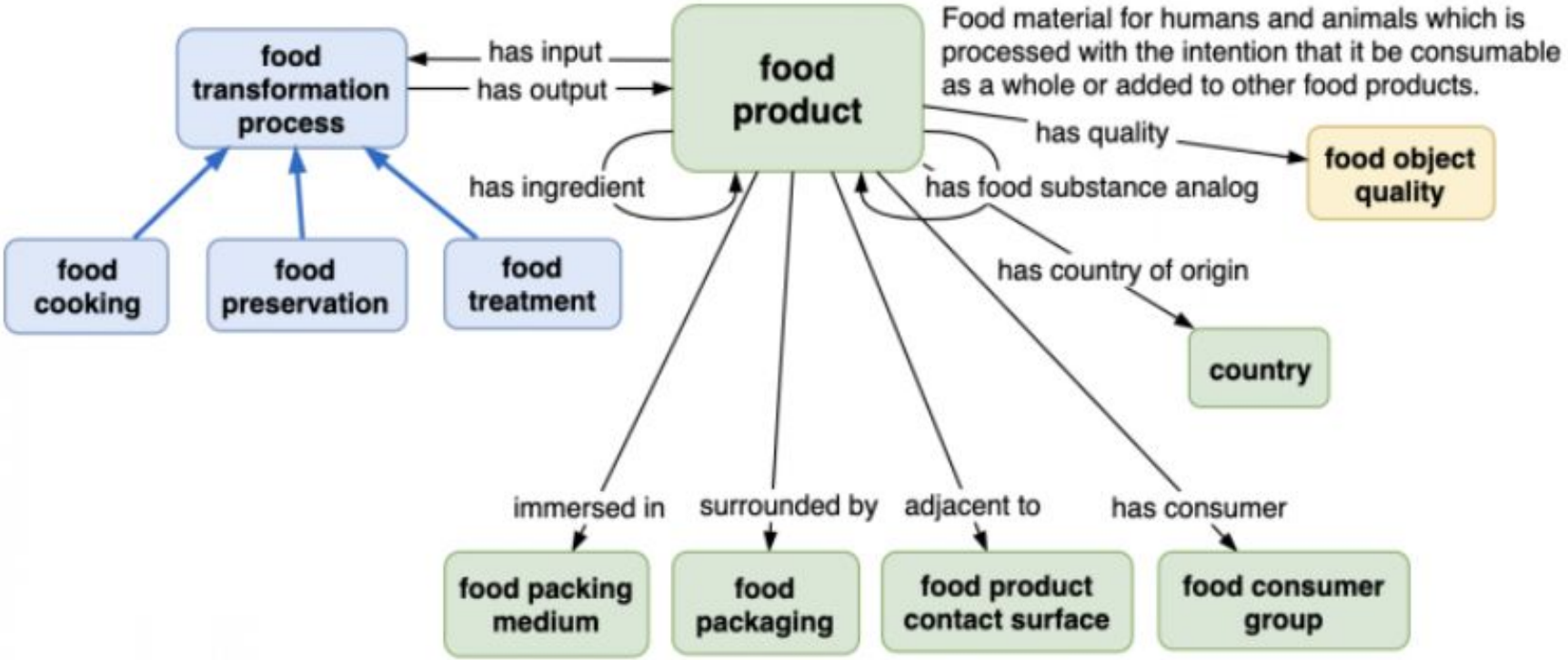
H6 | Quantitatie experiment: own dataset - 74 rules for substitutions (pairs: ingredeint & a few substtutes); 3 methods compared, first-acc@1, @5; secondAcc@1, @5; Qualitative experiment: visualization - embeddings cloud for regular

	A	D	E	F	G	H
1	reference	domain (e-commerce /food/linguistics...)	approach - method etc.	dataset - name, features etc.	input representation (embeddings)	evaluation - metrics, results, etc.
23	Achananuparp, Palakorn, and Ingmar Weber. "Extracting Food Substitutes From Food Diary via Distributional Similarity." ArXiv:1607.08807 [Cs], July 29, 2016. http://arxiv.org/abs/1607.08807.	food	PPMI (positive pointwise mutual information), SVD, dot product as the similarity measure. Food products are represented using their main ingredients from a fixed taxonomy.	Webscrapping of MyFitnessPal, enhanced with a manually built taxonomy (I. Weber and P. Achananuparp. Insights frommachine-learned diet success prediction. InProceedingsof Pacific Symposium on Biocomputing (PSB), 2016.)	SVD	MAP, NDCG, precision@1, precision@10
24	Akkoyunlu et al., "Investigating Substitutability of Food Items in Consumption Data."	food	A meal is represented as an itemset of food products, each meal is a node in a graph, nodes are connected iff they differ in at most one product. On the graph maximal cliques are mined and filtered, so each represents a context + a set of substitutable items. The set is then ranked according to a heuristic score inspired by the Jaccard index.	INCA 2 - the result of a 2006/7 survey on food consumption. ~4000 respondents. Possibly in French.	itemsets	none
25	Gaillard, Lieber, and Nauer, "Improving Ingredient Substitution Using Formal Concept Analysis and Adaptation of Ingredient Quantities with Mixed Linear Optimization."	food	Formal Concept Analysis (FCA) + heuristic search otherwise. Mixed linear optimization to adapt quantities.	1. WikiTaaable knowledge base, http://wikitaaable.loria.fr/, but the URL doesn;t seem to be alive anymore; 2. Crawled from yummly.com, 1327 recipes.	itemsets	none

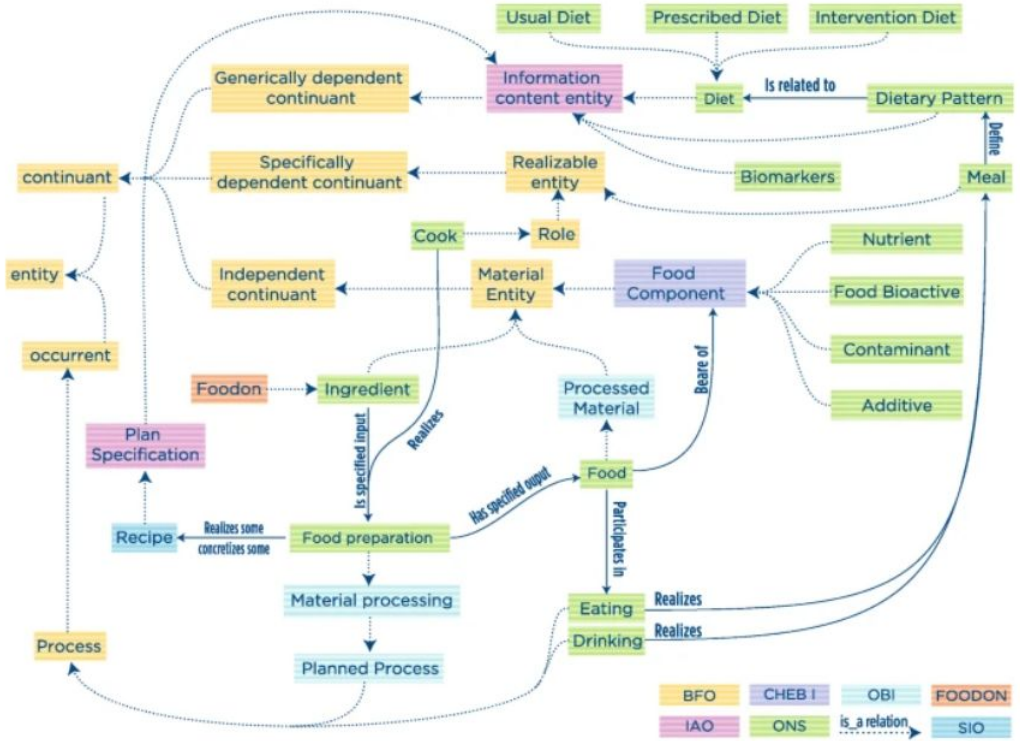
Ontologies in the Food Domain



FoodOn - facets of food, from harvest to consumption

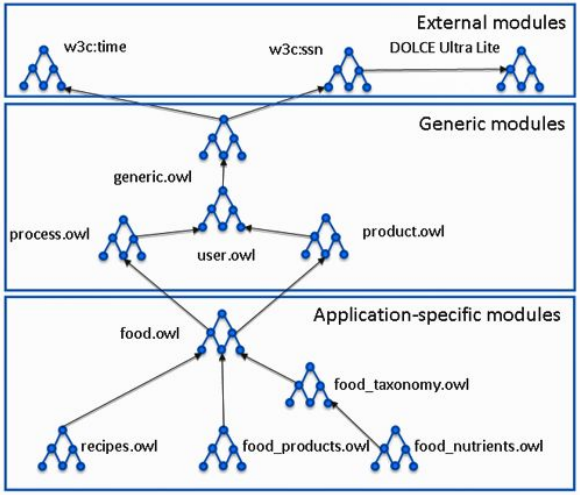


Ontology for Nutritional Studies



Other ontologies in the food domain

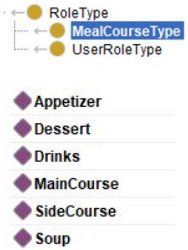
- **FoodKG**
- **Ontology for Nutritional Studies (ONS)**
- **The SmartProducts Network of Ontologies (SPO)**



Comparative analysis of the ontologies

Selected aspect	ONS	FoodKG	SPO	FoodOn
Class defining a recipe	✓	✓	✓	✓
Class defining a meal	✓	✓	✗	✓
Class defining a component	✓	✓	✗	✗
Class defining a product substitute	✗	✗	✗	✓
Class defining food	✓	✓	✓	✓
Classes regarding drinking	✓	✗	✓	✓
Classes regarding processed food	✓	✓	✗	✓
Classes describing process of making food	✓	✗	✓	✓
Classes regarding diet	✓	✗	✓	✓
Classes regarding user preferences	✓	✗	✓	✓
Classes regarding product quality	✓	✓	✓	✓
Classes regarding taste	✓	✓	✗	✓

Meal

ONS	FoodKG	SPO	FoodOn
YES	**YES**	**NO**	**YES**
,,A meal is an eating/drinking occasion which usually occurs at certain time during the day. During a meal, through processes of eating and drinking, food are consumed''	„a class that links a recipe to a certain meal"	**'MealType', 'MealCourseType' i 'MealCourse-Role'.**	„Any of the occasions for eating food that occur by custom or habit at more or less fixed times."
			'meal (ready-to-consume)' 'meal replacement (us cfr)' 'meal replacement (weight-reducing)' 'meal replacement food product' 'meal replacement food'

Recipe

ONS	FoodKG	SPO	FoodOn
YES	YES	YES	YES
has_part some 'action specification' has_part some 'objective specification' 'is about' some entity 'is about' some 'realizable entity' **has_part some 'action specification'** **has_part some 'objective specification'** **'is about' some entity** **'is about' some 'realizable entity'**	recipe breakfast recipe dinner recipe high glycemic recipe less than one hour recipe lunch recipe side recipe **'almond biscotti'** **'baked chicken tender'** **'banana blueberry almond flour muffin'** **'banana bread'** **'beef nilaga'**	**hasIngredient min 1 IngredientPortion** **hasIngredient only IngredientPortion** **hasStep only CookingActivity** **WorkflowDefinition**	plan specification **device specification** **food recipe** **ingredient specification** **step specification**

Substitution

Ontology	Substitution	Context	Limitations
FoodOn	a (symmetric) relation: 'has food substance analog'	dietary and allergen analysis	no links with food preparation process, recipes
	subclasses of class: 'food product analog'		
FoodKG	heuristics based on explicit semantics and embeddings	dietary restrictions nutritional change	no ontological conceptualization
ONE	no term(s) for substitution		
ONS	no term(s) for substitution		

Knowledge modelling in the food domain

1. Conceptualizing what substitution is

Three aspects of substitution

There is no such phenomenon as substitution per se. Substitution:

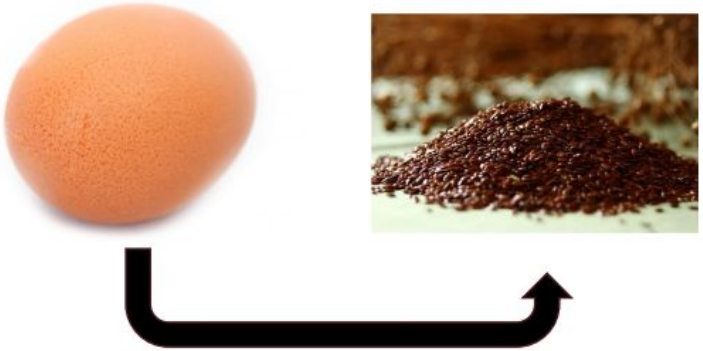
1. *is always made in a specific situation (**context**),*
2. *has a direction and/or a specific goal (**direction**),*
3. *is done with similar or related to object A, object B (**similarity/relatedness**).*

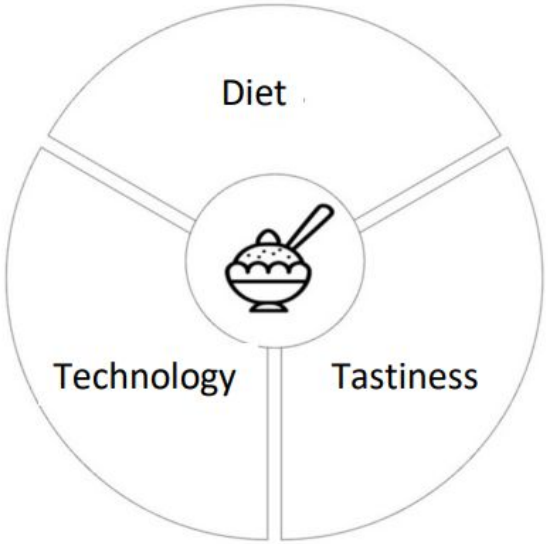
What ingredient to substitute, why and how?

Context:

Constraints: allergy, diet, health condition, lack of ingredient, …

Goals: exclude a specific product for health reasons, increase the intake of a particular nutrient, make a dish less dense, ...



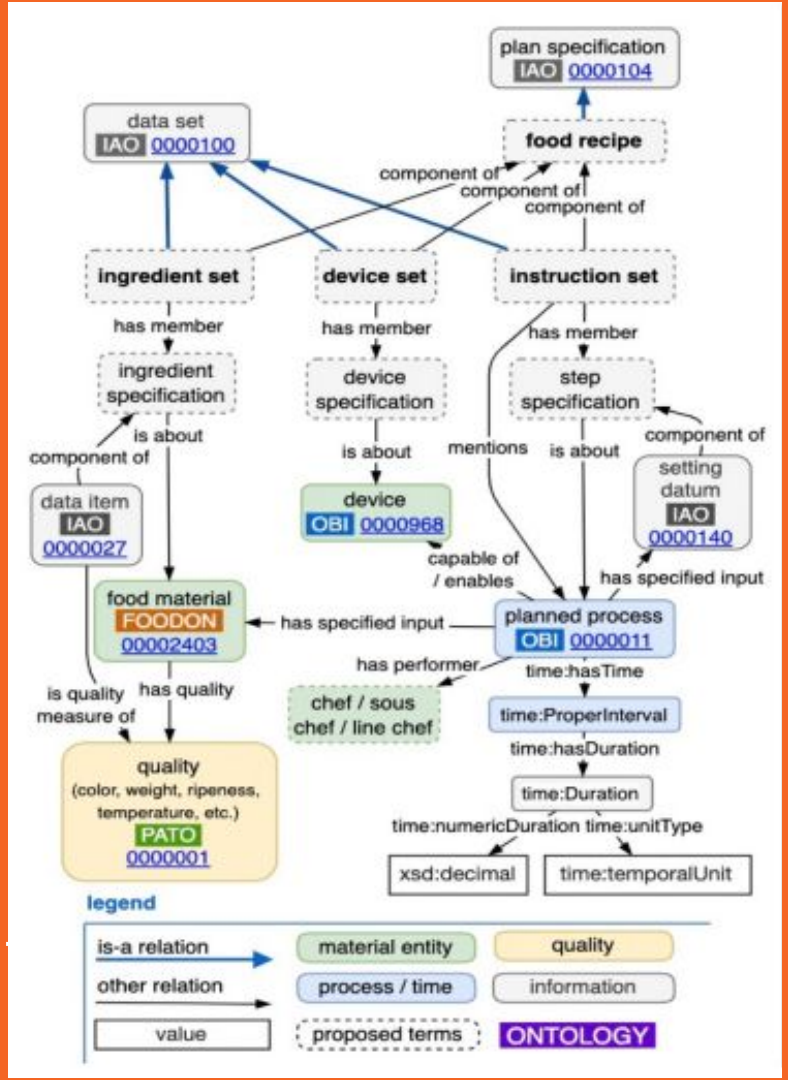


Modeling substitution

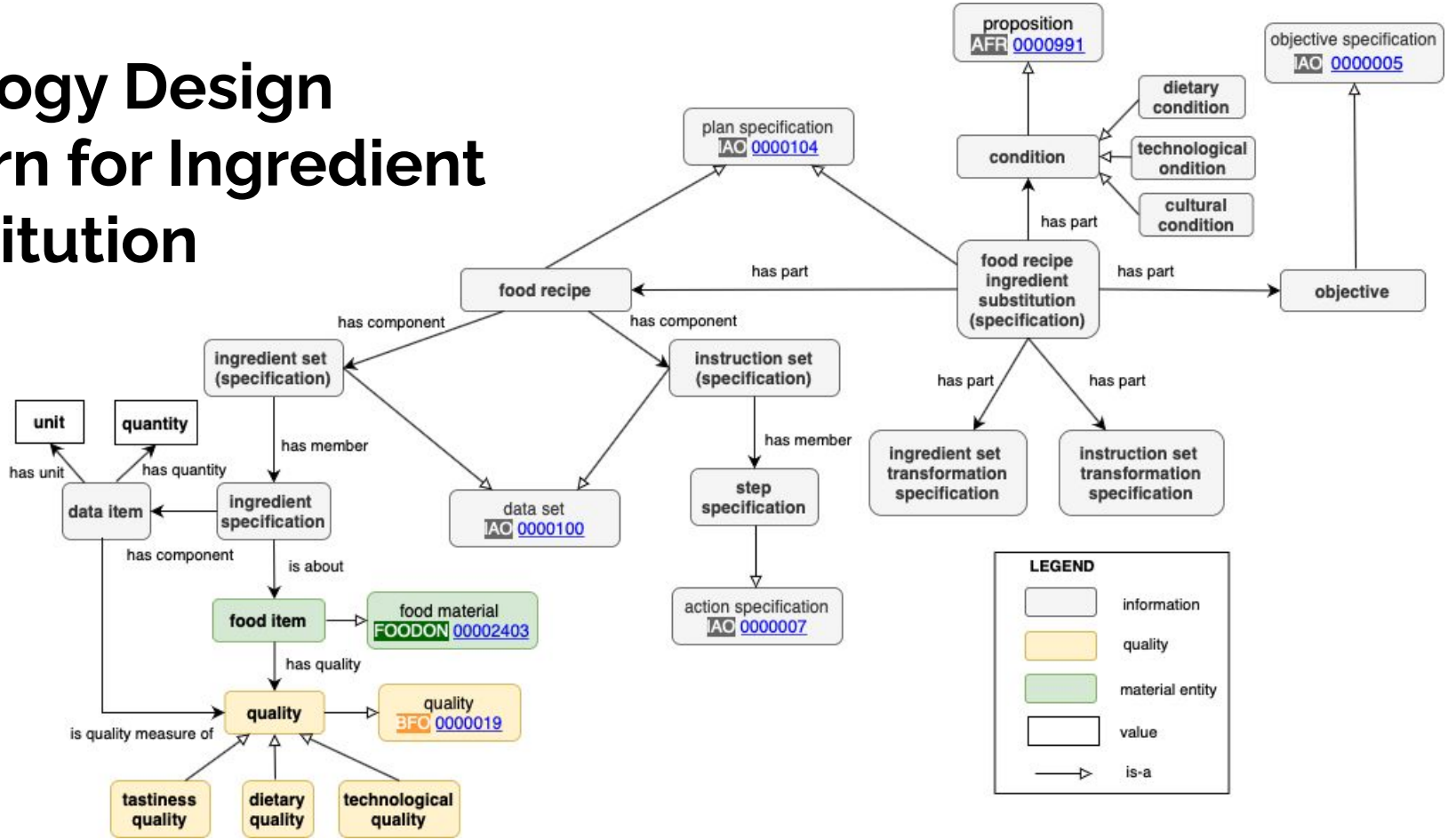
- What related concepts should we include (context)?
- How to link our new model with the existing ones?
- Substitution: binary (simple) vs. n-ary relation
- What is a recipe: is it a plan, a process, a specification?

Recent recipe ODP adopted in FoodOn

subclass of plan specification
ingredients + instructions + devices
links to FoodOn (food material) and other ontologies



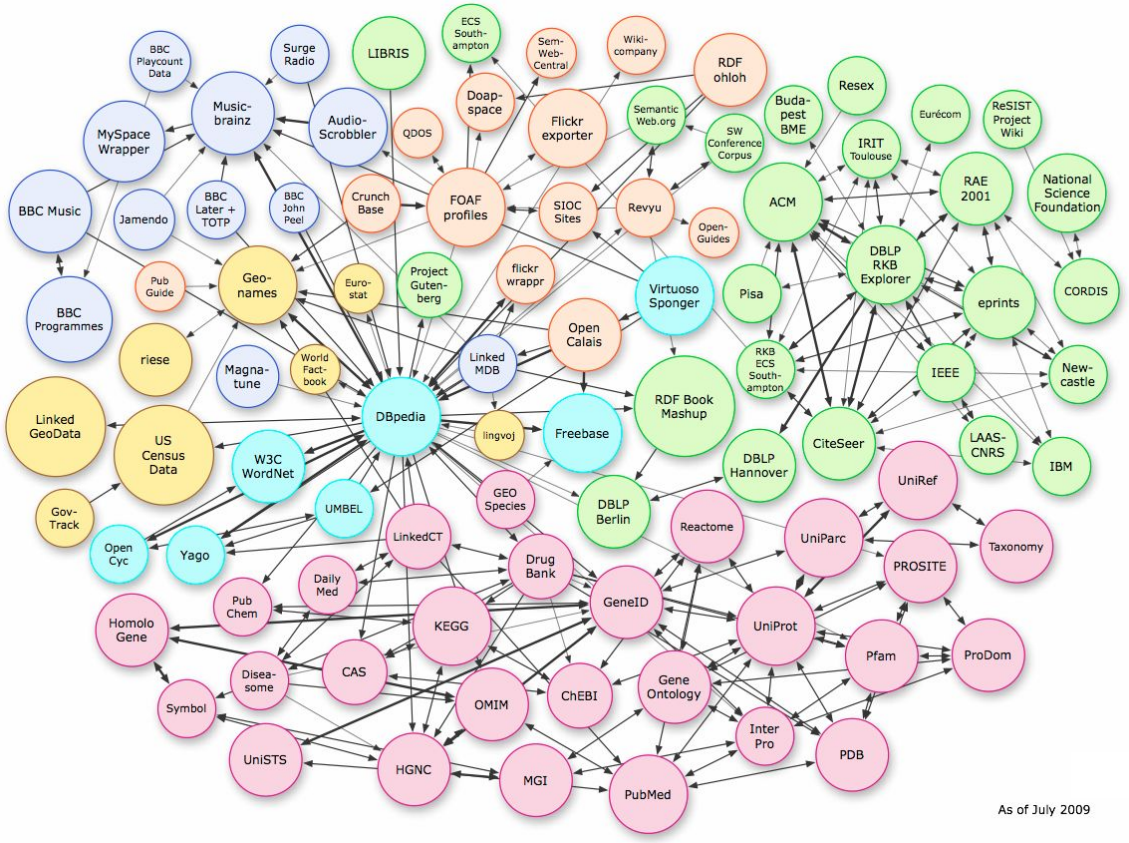
Ontology Design Pattern for Ingredient Substitution



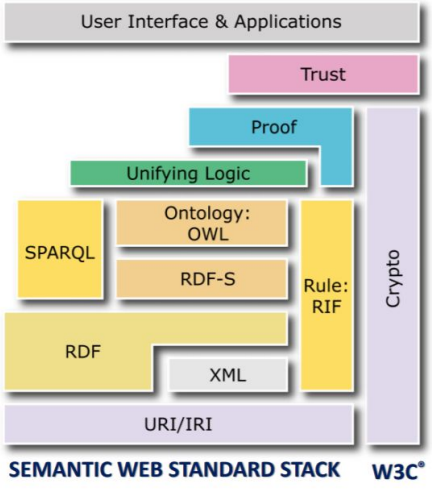
Ławrynowicz, A., Wróblewska, A., Adrian, W.T., Kulczyński, B. and Gramza-Michałowska, A., 2022. **Food Recipe Ingredient Substitution Ontology Design Pattern.** ***Sensors*****, *22*(3), p.1095.**

2. Building a knowledge graph for ingredient substitution

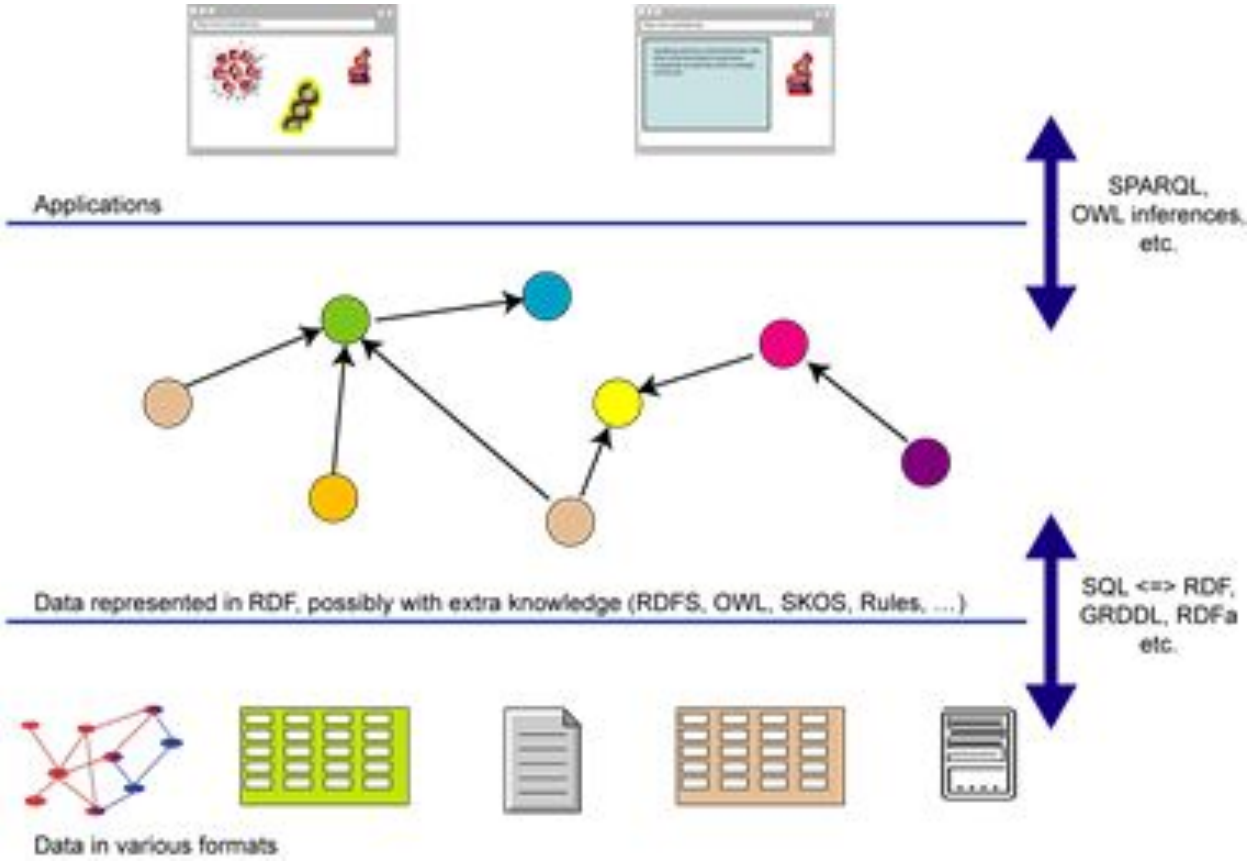
Once upon a time…



As of July 2009

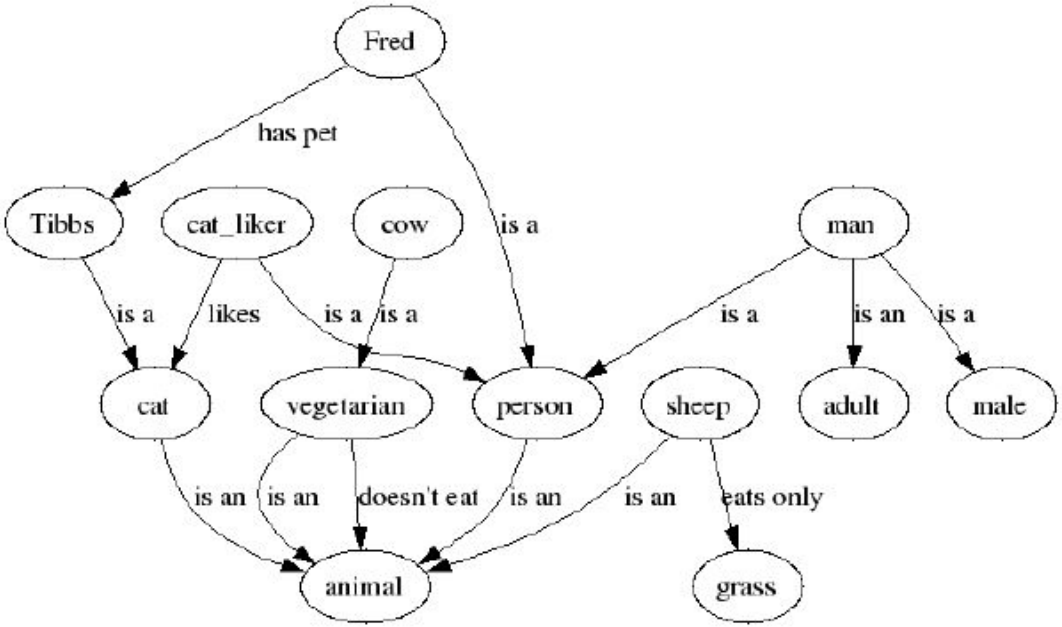


1: abstract representation layer based on graphs



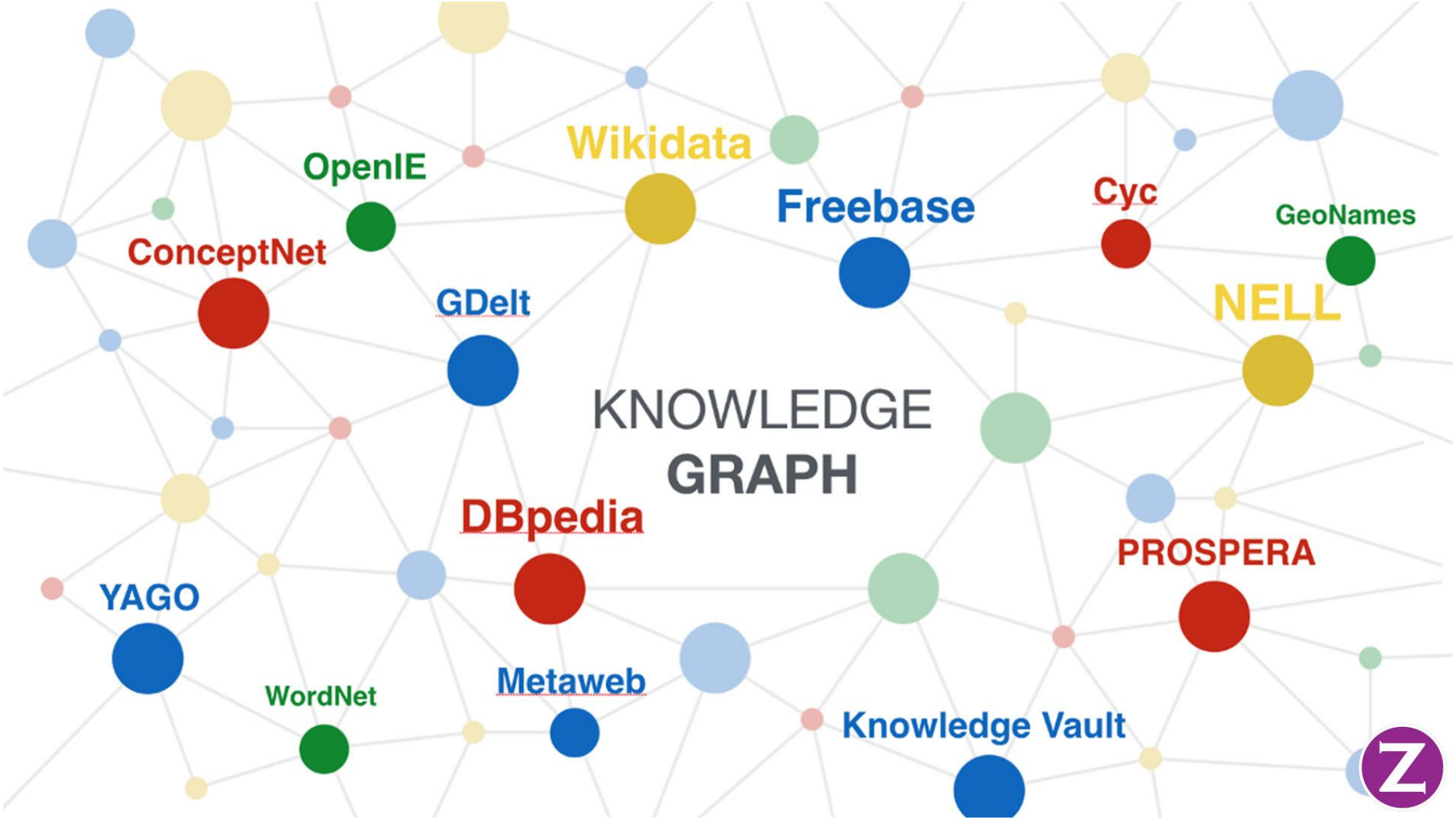
- $Man \equiv Person \sqcap Adult \sqcap Male$,
- $CatLiker \sqsubseteq \exists likes.Cat$, $Sheep \sqsubseteq \forall eats.Grass$,
- $Person(fred)$, $Cat(tibbs)$, $hasPet(fred, tibbs)$

2: ontologies: formal specification of a conceptualization



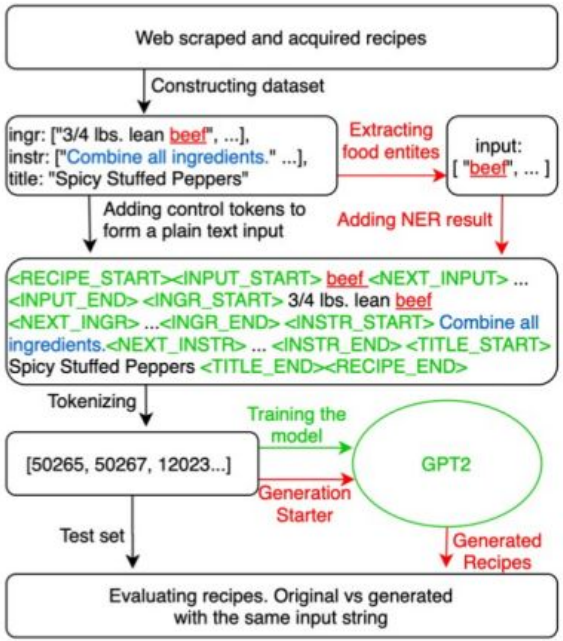
Semantic Web 20 years later...

Google introduces Knowledge Graph…


Wikidata
OpenIE
Freebase
Cyc
GeoNames
ConceptNet
GDelt
NELL
KNOWLEDGE
GRAPH
DBpedia
PROSPERA
YAGO
Metaweb
WordNet
Knowledge Vault

Data: RecipeNLG

RecipeNLG dataset, partly based on the Recipe1M+ dataset, and providing over 1 million new, preprocessed and deduplicated recipes



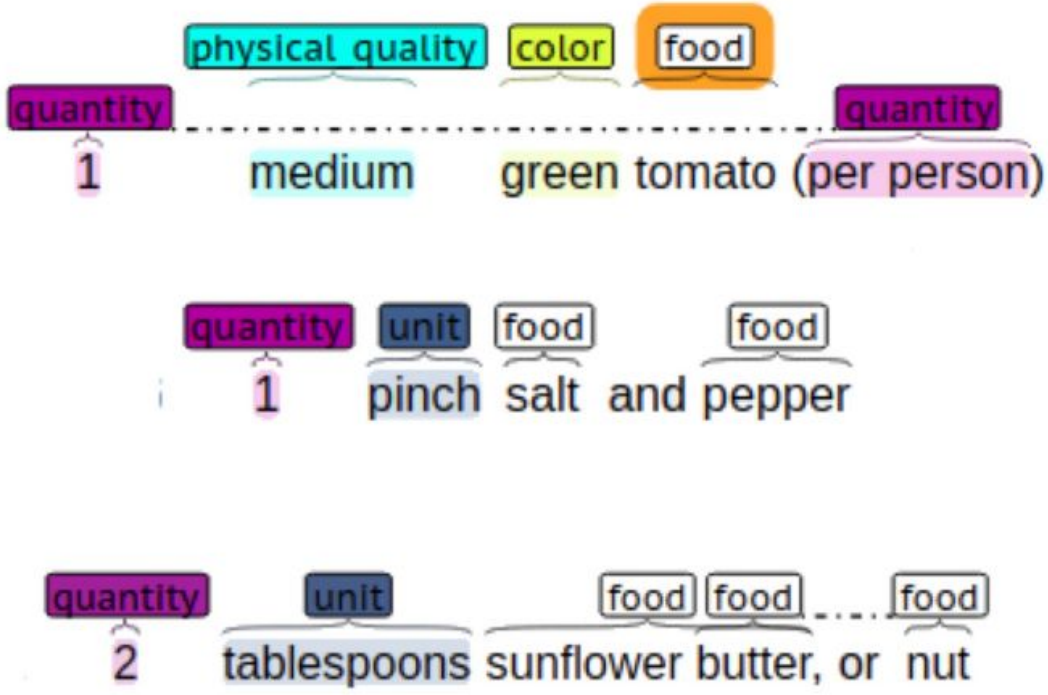
Michal Bien, Michal Gilski, Martyna Maciejewska, Wojciech Taisner, Dawid Wisniewski, Agnieszka Lawrynowicz, RecipeNLG: A Cooking Recipes Dataset for Semi-Structured Text Generation. INLG 2020: 22-28

TASTEset - Recipe Dataset and Food Entities Recognition Benchmark

Tagset:

- FOOD as the name of an ingredient (e.g. bread, mayonnaise, salt, tomato),
- QUANTITY as a quantity (usually expressed by digits or a float),
- UNIT as a unit of measurement (e.g. bunch, cup, grams, jar, millimeters, slices, stalks, tablespoon, teaspoon)
- PROCESS as the attribute of the ingredient, usually referring to an action to be taken to prepare the ingredient (e.g. chopped for parsley, crushed for garlic, grated for ginger, minced for garlic),
- PHYSICAL QUALITY as the characteristic of the ingredient (e.g. boneless for chicken breast),
- COLOR as the color of the ingredient,
- TASTE as the flavour (e.g. bittersweet, butter-flavoured, sweet, semi-sweet),
- PURPOSE as the purpose of using the ingredient in the recipe (e.g. for dusting about flour, for garnish about sunflower seeds, for frying about canola oil, and as topping about sour cream),
- PART as a part of the ingredient required by the recipe (e.g. yolks and whites as parts of the eggs).

TASTEset: entity examples

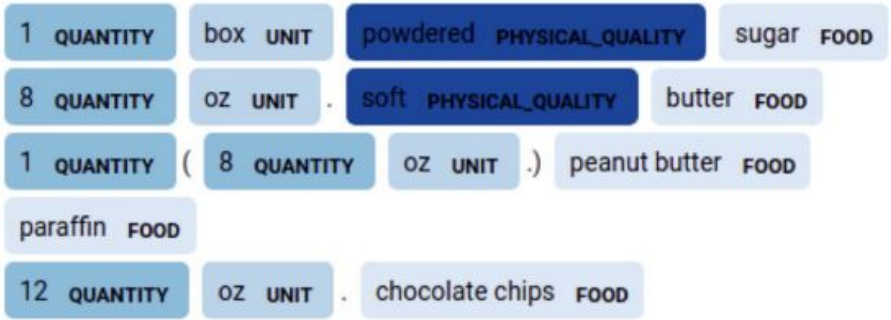


TASTEset

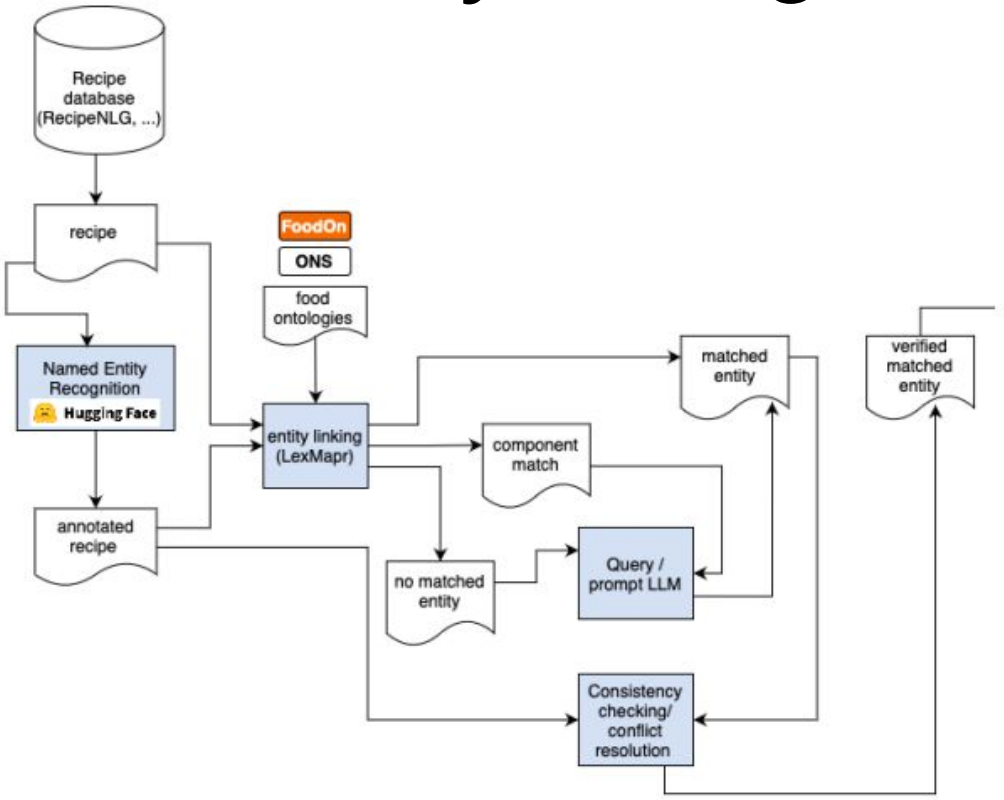
- 700 recipes with more than 13,000 entities to extract
- manual annotation covered 3,788 ingredients of varying complexity
- state-of-the-art baselines of named entity recognition models

List of ingredients

- 1 box powdered sugar
- 8 oz. soft butter
- 1 (8 oz.) peanut butter
- paraffin
- 12 oz. chocolate chips



Named Entity Recognition + Entity Linking

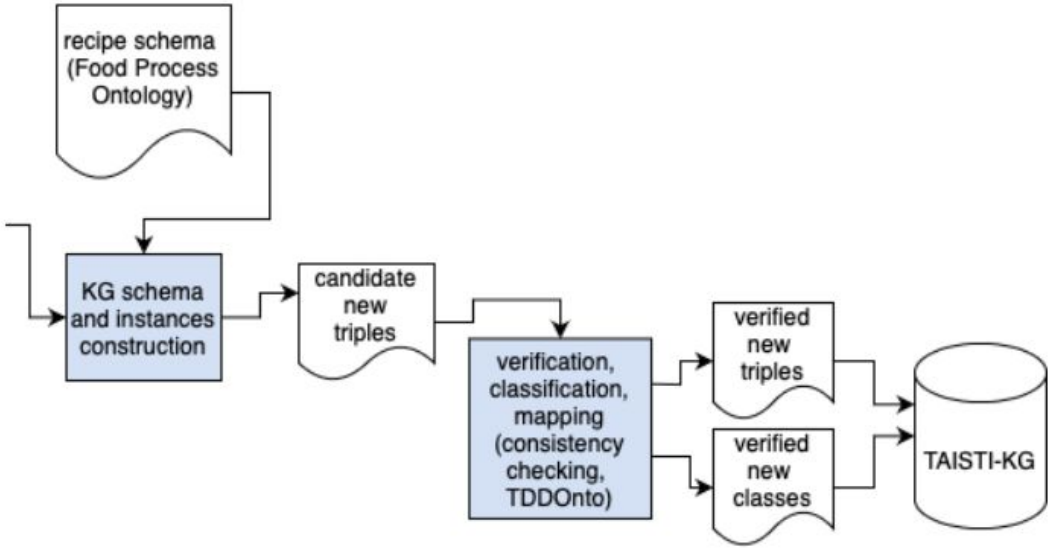


Entity linking - LexMapr

green onions (chopped)

['onion (chopped):FOODON_03316397', 'scallion (whole, raw):FOODON_03311340'] Component Match

Knowledge graph authoring



Graph models and frameworks

Property graph

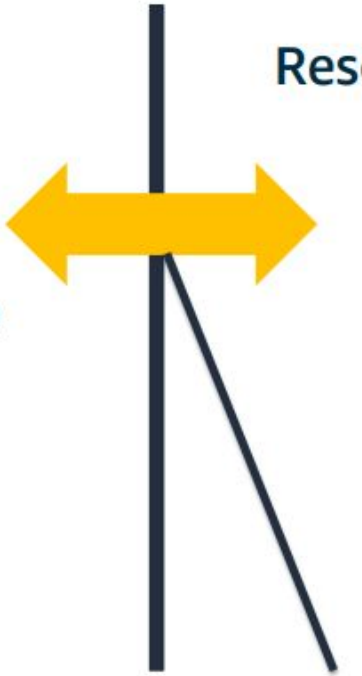
Apache TinkerPop Gremlin

openCypher

Property Graph Query Language (PGQL)

GQL

Others



Resource Description Framework (RDF)

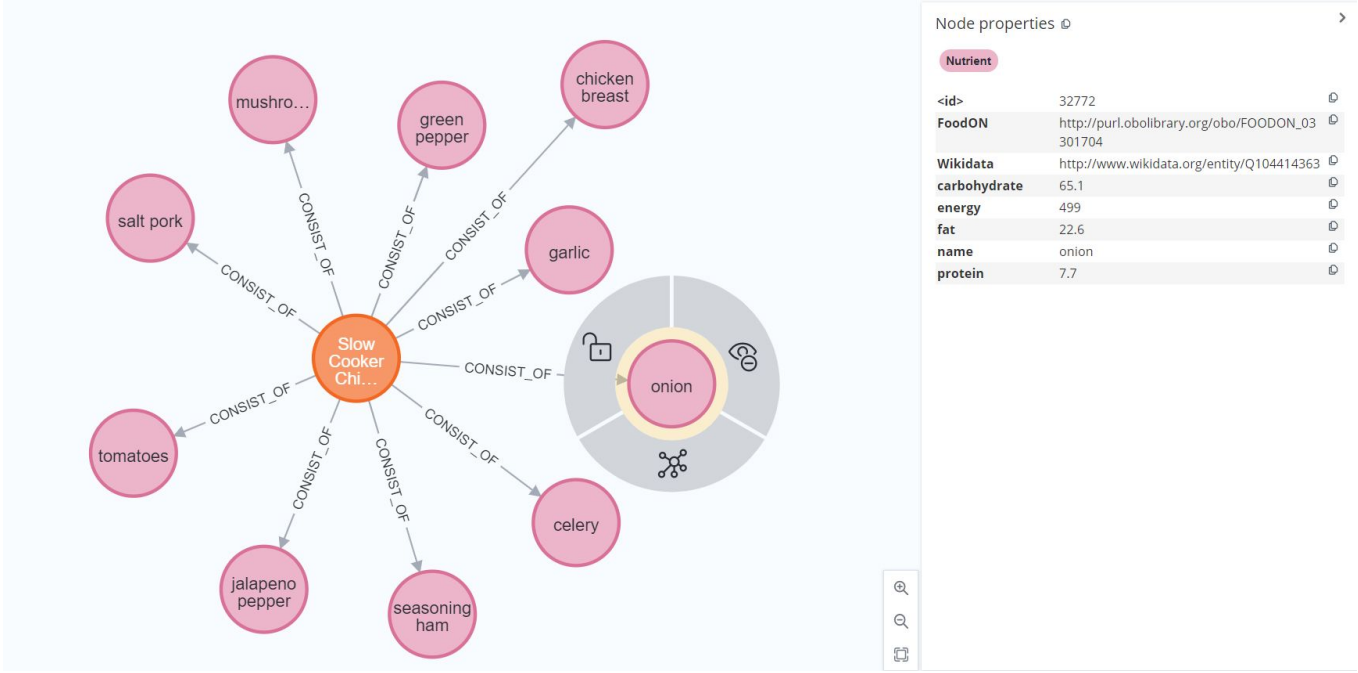
W3C standard

SPARQL Query Language

W3C RDF

W3C SPARQL

“Knowledge graph” in neo4j



“Knowledge graph” in neo4j

Turkey Mushr…

Node properties

Recipe

<id>	21
instructions	1.Heat the oil in a skillet over medium heat, and cook the turkey until evenly browned. Season with garlic powder, Italian seasoning, and pepper. Drai… Show all
name	Turkey Mushroom Stew

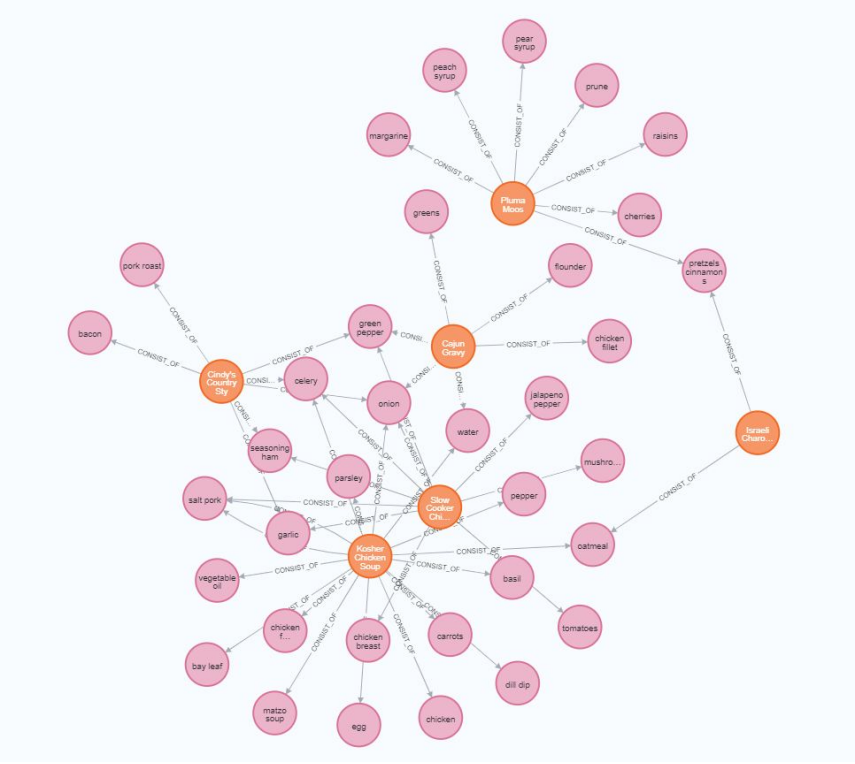
apple

Node properties

Nutrient

<id>	29955
FoodON	http://purl.obolibrary.org/obo/FOODON_03310788
Wikidata	http://www.wikidata.org/entity/Q65524294
carbohydrate	65.9
energy	243
fat	0.32
name	apple
protein	0.93

“Knowledge graph” in neo4j



▼ 0 {6}
name : milk
▼ Protein {2}
Amount : 1.03
unit : g
▼ Energy {2}
Amount : 70
unit : kcal
▼ Carbohydrate, by difference {2}
Amount : 6.89
unit : g
▼ Total lipid (fat) {2}
Amount : 4.38
unit : g
▼ foodPortions [3]
▼ 0 {2}
gramWeight : 246
portionDescription : 1 cup
▼ 1 {2}
gramWeight : 30.8
portionDescription : 1 fl oz
▼ 2 {2}
gramWeight : 0
portionDescription : Quantity not specified

Reasoning about the integrated knowledge

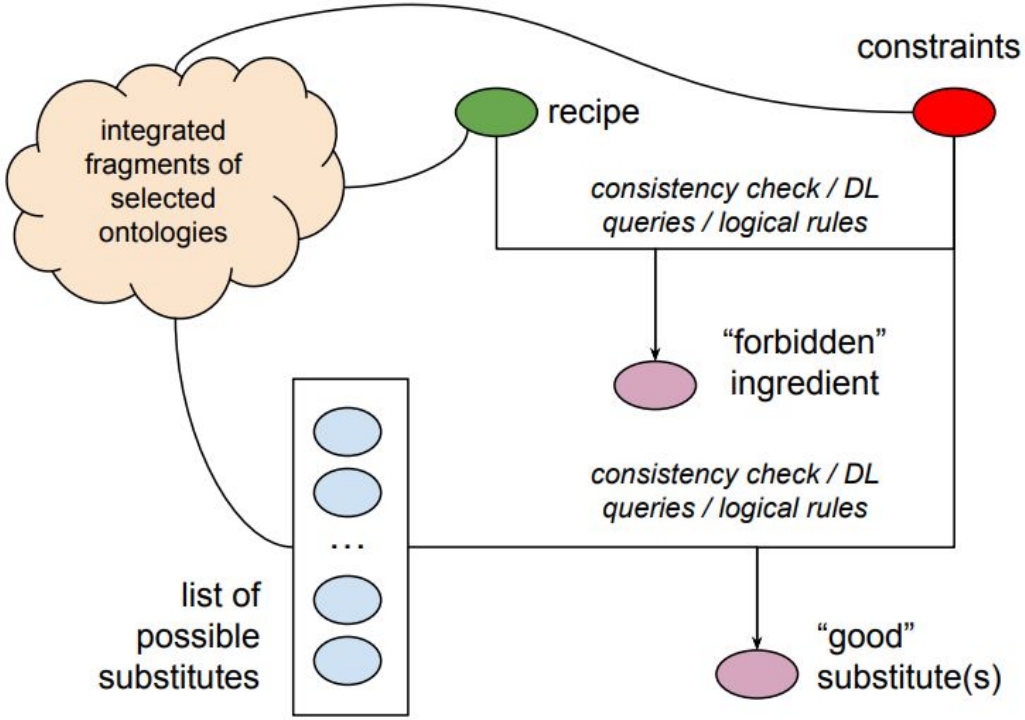
Research objectives for reasoning in TAISTI (not everything!)

1. To propose a *logic-based method* for **identifying ingredients** in food recipes that do not meet the specified *constraints*
2. To propose a logic-based method for **pruning "wrong" substitutes** from a list of substitutes proposed by ML models

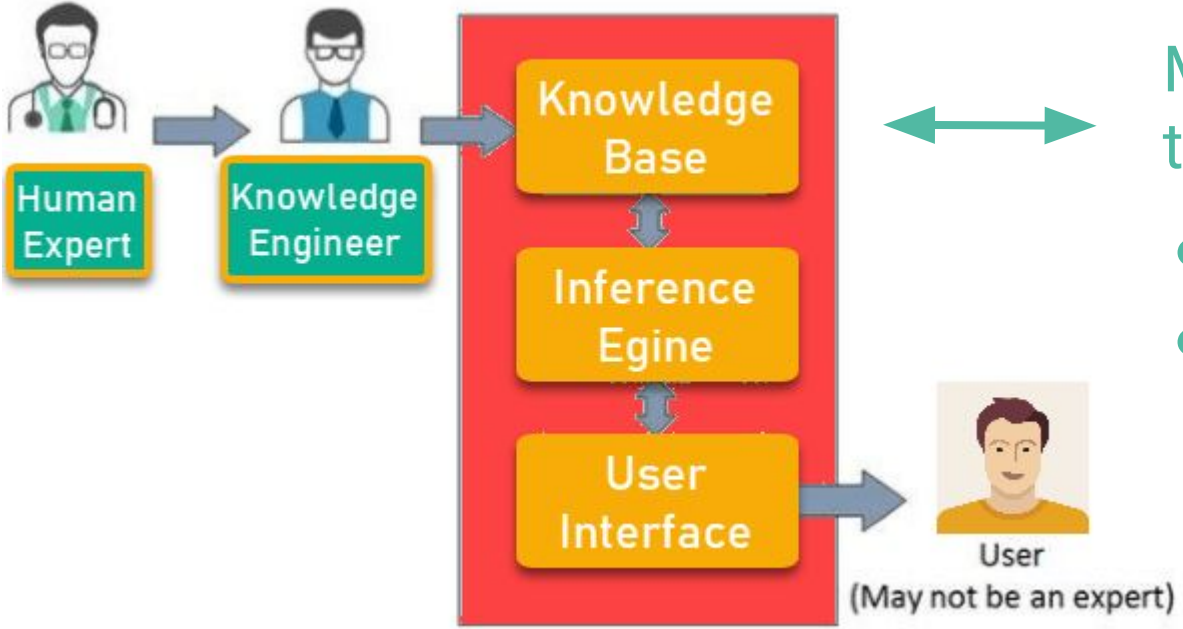
In fact, more questions arise from the above….

- What knowledge do we need?
- How to model constraints (and everything else)?
- How to reason about this knowledge?

Overview of the proposed model



Methodology: as for a knowledge-base system



Main operations on the knowledge base:

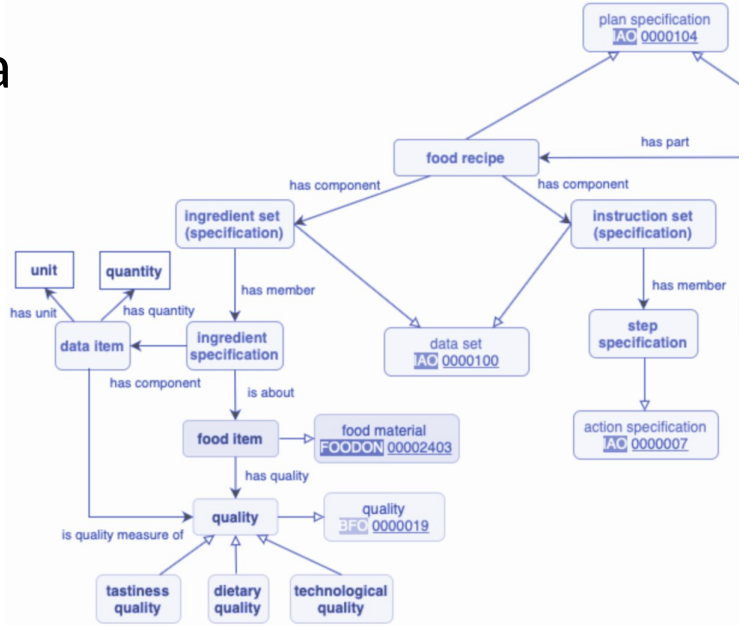
- TELL
- ASK

Image source: https://www.guru99.com/expert-systems-with-applications.html


TELL - what knowledge do we need to put in the KB?

1: Recipe datasets → knowledge graph

- Semi-structured data sources
  - text, images
  - ingredients sets
  - instructions sets
  - nutritional data
- Data acquisition
  - APIs
  - crawling, scrapping
  - database dumps
- Annotation + Entity linking



2. Food-related ontologies and knowledge graphs

- FoodOn - “A farm to fork ontology”
- Ontology for Nutritional Studies, Ontology of Nutritional Epidemiology
- Smart Products Ontology
- FoodKG - knowledge graphs with recipes
- Tables of nutritional data
- EU regulations on allergies vocabulary

EVERYTHING IS (OR SHOULD BE) CONNECTED…

3. Domain experts' knowledge

- How domain experts talk:
  - "It's not that easy…"
  - **"It depends…"**
  - "It works for baking, but not for frying…"
  - "Everything could be substituted, the possibilities are endless…"
  - "Why would one want to substitute it?"

Knowledge graph/ontology design

Classes:

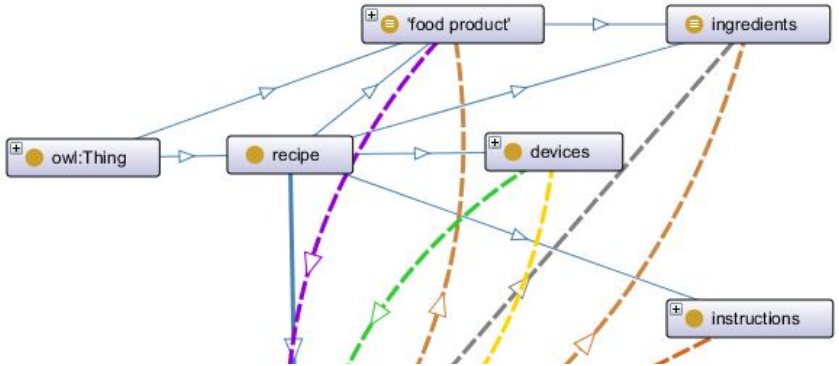
- Recipe
- Ingredient spec.
- Food product
- **Diet**
- **Allergy**

Object properties:

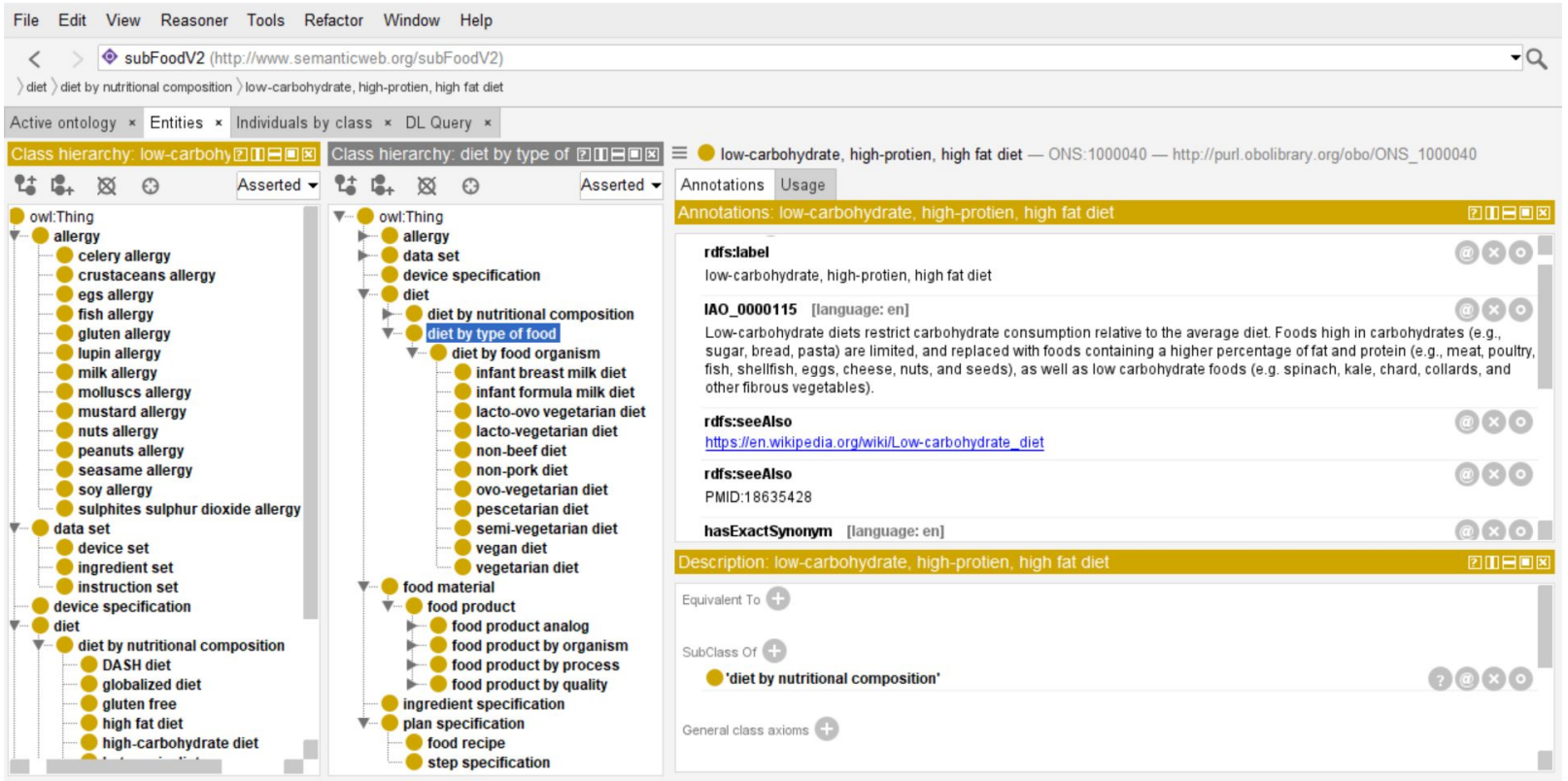
- **acceptableIn** - (food product → diet)
- **unacceptableIn** - food product/recipe → diet
- **hasAllergicTrigger** - (food product → allergy)
- hasIngredient/isNeededFor - (food product ←→ recipe name)
- **isSubstituteFor** - food product → food product
- useFor - instructions (specific steps) → recipe name

Data properties:

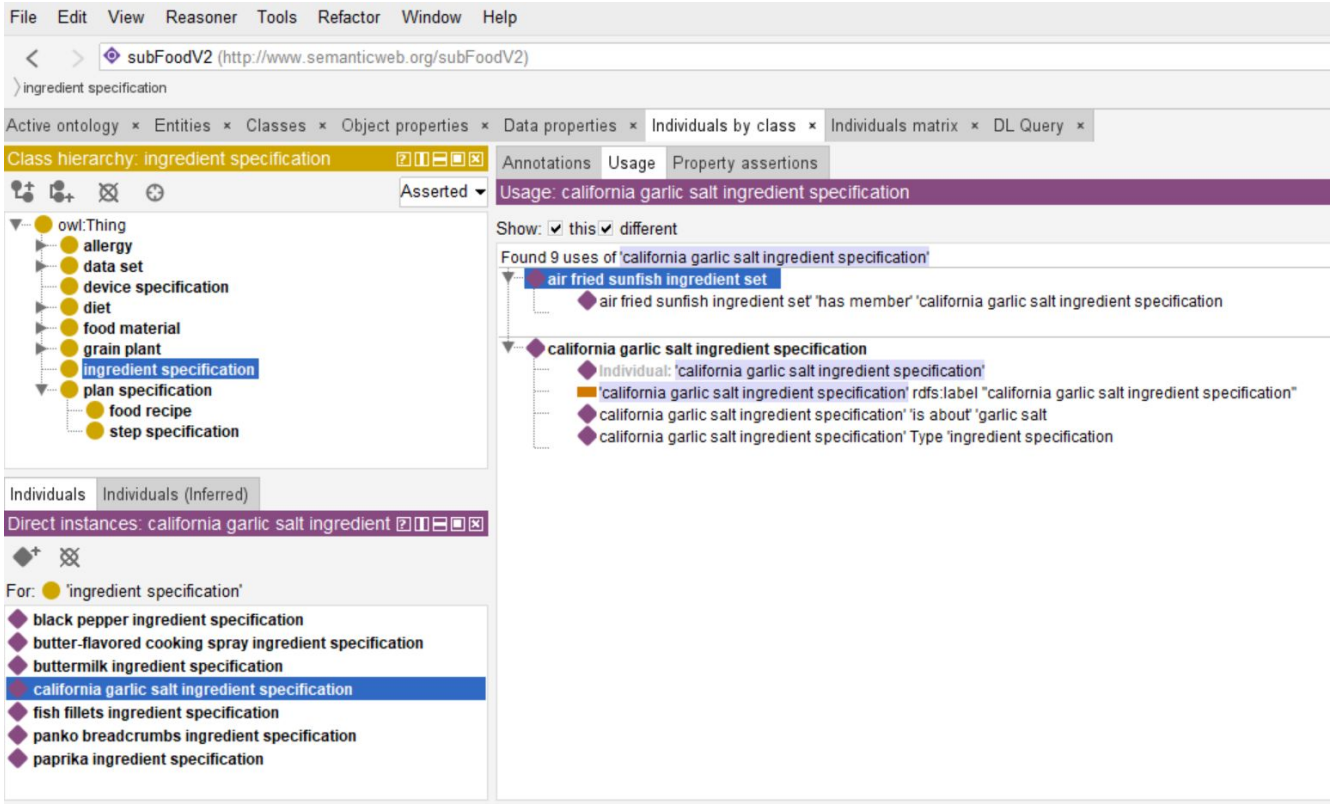
- *hasCalorificContentValue*
- *hasCarbohydratesContentValue*
- *hasFatContentValue*
- *hasProteinContentValue*
- *hasVitaminsContentValue*



Ontology imports and populating with instances



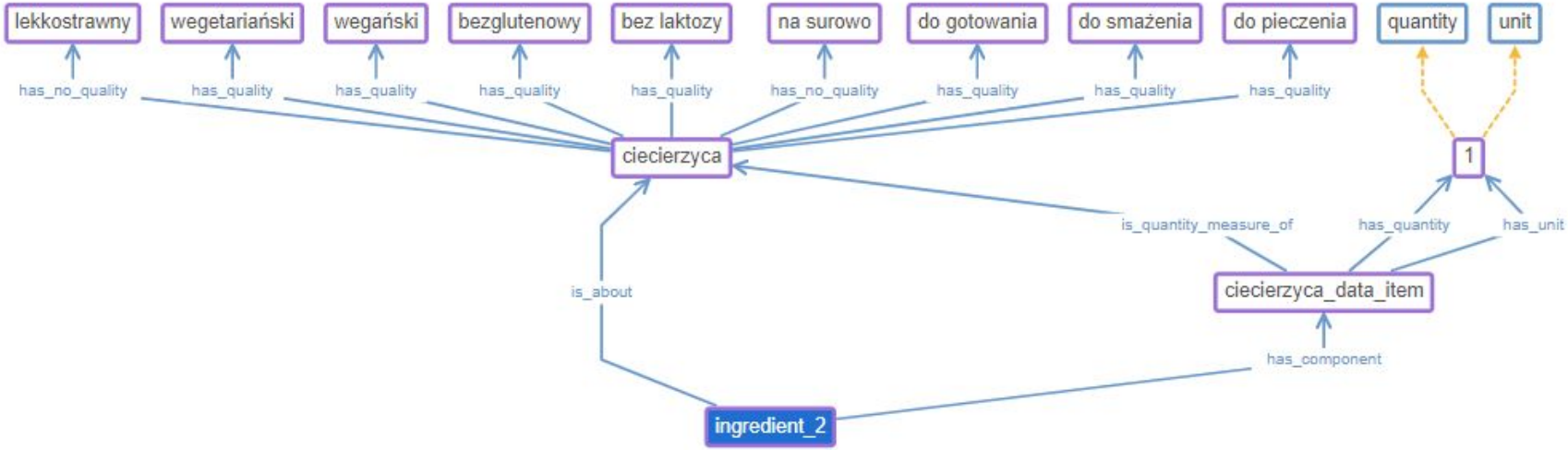
Ontology imports and populating with instances



Instantiation of the substitution pattern

(experts modeling the actual substitution)

KARTA PRODUKTU DLA CIECIERZYCY				
Etykiety ogólne				
lekkostrawny	**bezglutenowy**	**wegetariański**	**wegański**	**bez laktozy**
NIE	TAK	TAK	TAK	TAK
Etykiety technologiczne				
na surowo	**do gotowania**	**do smażenia**	**do pieczenia**	
NIE	TAK	TAK	TAK	
Etykiety odżywcze				
źródło	**mało**	**kalorie**		
białko, węglowodany	witaminy	120		




ASK - what questions
to ask and how to
answer them?

Ontology-based query answering

,,With what vegan product can we substitute beef in a ‘roast romaine’ recipe?”

```
SELECT  DISTINCT ?recipe ?food_item
    WHERE{ ?recipe_ingr_sub rdf:type food:food_recipe_ingr_subst_spec.
        ?recipe rdf:type food:food_recipe.
        ?set rdf:type food:ingredient_set_specification.
        ?spec rdf:type food:ingredient_specification.
        ?spec2 rdf:type food:ingredient_specification.
        ?food_item rdf:type food:food_item.
        ?set_trans rdf:type food:ingredient_set_trans_spec.

        ?recipe_ingr_sub food:has_part ?recipe.
        ?recipe_ingr_sub food:has_part food:pieczen_rzymska.
        food:pieczen_rzymska food:has_component ?set.
        ?set food:has_member ?spec.
        ?spec food:is_about food:wolowina.
        ?recipe_ingr_sub food:has_part ?set_trans.
        ?set_trans food:has_member ?spec2.
        ?spec2 food:is_about ?food_item.
        ?food_item food:has_quality food:weganski.
    }
```

Ontological reasoning

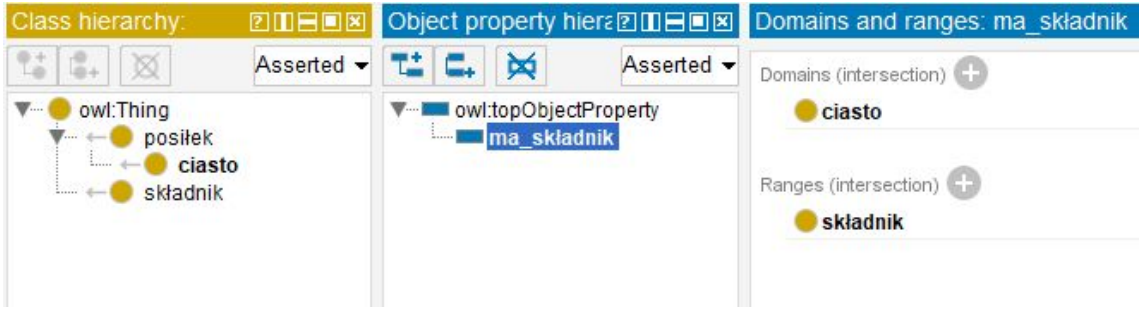
Individuals by type: wieprzowina

- 'food_recipe_ingredient_substitution_(specification)' (1)
- 'ingredient_set_(specification)' (1)
- 'instruction_set_(specification)' (1)
- condition (1)
- data_item (2)
- dietary_quality (5)
- food_item (11)
- food_recipe (1)
- ingredient_set_transformation_specification (1)
- ingredient_specification (2)
- instruction_set_transformation_specification (1)
- quality (1)
- quantity (2)
- step_specification (5)
- technological_quality (4)
- unit (2)
- wieprzowina

Individuals by type (inferred): wieprzowina

- 'food_recipe_ingredient_substitution_(specification)' (1)
- 'ingredient_set_(specification)' (1)
- 'instruction_set_(specification)' (1)
- condition (1)
- data_item (2)
- dietary_quality (5)
- food_item (12)
  - bulka_tarta
  - wolowina
  - cebula
  - cukier_brazowy
  - musztarda
  - wieprzowina
  - ciecierzyca
  - mleko
  - keczup
  - jajo
  - sol
  - pieprz
- food_recipe (1)
- ingredient_set_transformation_specification (1)

Ontological reasoning



Modeling rules and DL queries (1a)

High-level queries for identifying unacceptable or even dangerous ingredients

For ***recipe*:Recipe** and ***my_diet*:Diet,** we obtain an unacceptable ingredient ***a***

```
FoodProduct(a) AND Recipe(recipe) AND Diet(my_diet) AND
isNeedeFor(a, recipe) AND unacceptableIn(a, my_diet)
```

For ***recipe*:Recipe** and ***my_allergy*:Allergy,** we get a dangerous ingredient ***a***

```
FoodProduct(a) AND Recipe(recipe) AND
Allergy(my_allergy) AND isNeedeFor(a, recipe) AND
hasAllergicTrigger(a, my_allergy)
```

Modeling rules and DL queries (2a)

High-level queries for pruning “wrong” substitutes

For a given ***recipe:Recipe, my_diet:Diet*** and ***isSubstitute(a,b)***

```
FoodProduct(a) AND Recipe(recipe) AND isNeededFor(a, recipe)
AND FoodProduct(b) AND isSubstituteFor(a,b)
AND Diet(my_diet) AND unacceptableIn(b, my_diet)
```

For a given ***recipe:Recipe*** and ***my_allergy:Allergy***

```
FoodProduct(a) AND Recipe(recipe) AND isNeededFor(a, recipe)
AND FoodProduct(b) AND isSubstituteFor(a,b)
AND Diet(my_diet) AND hasAllergicTrigger(b, my_allergy)
```

Logic model in Answer Set Programming

Answer Set Programming

- Declarative programming paradigm
- Non-monotonic reasoning and logic programming
- Stable model semantics

Expressive KR language

- Roots in Datalog
- Default negation, disjunction, constraints, aggregates
- Weak constraints, functions, lists, sets, exist.quantifiers

Hard and weak constraints in ASP

Absolutely forbidden ingredients:

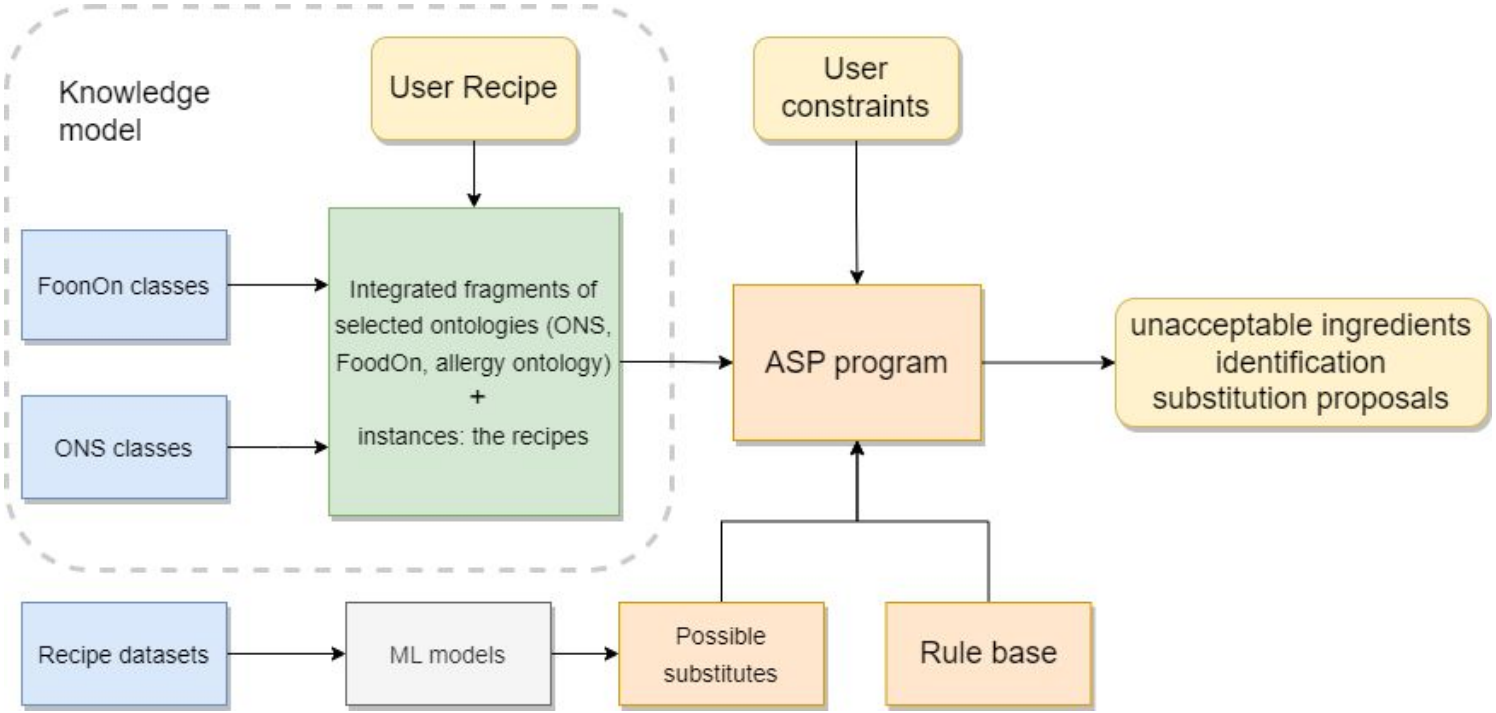
```
:- foodProduct(X), recipe(Recipe), allergy(my_allergy),
isNeededFor(X, Recipe), hasAllergicTrigger(X, My_allergy).
```

“If possible, avoid…” / “Optimize quantity of…”:

```
:~ foodProduct(X), recipe(Recipe),  diet(My_diet),
isNeededFor(X, Recipe), unacceptableIn(X, My_diet).  [1]
```

```
:~ foodProduct(Y), recipe(Recipe),  diet(My_diet),
isNeededFor(Y, Recipe), unacceptableIn(Y, My_diet).  [2]
```

Knowledge model (architecture of the solution)



Weronika T. Adrian, Maciej Kutyła, Jędrzej Potoniec, Agnieszka Ławrynowicz:
Reasoning about Ingredient Substitutions with Ontologies and Rules. RuleML+RR (Companion) 2023

Rule base of diets and allergies

% D1: vegan diet

```
unacceptableIn(X, vegan_diet) :- isSubclassOf(X, meat).
```

% D2: gluten-free diet

```
unacceptableIn(X,glutenfree_diet) :-
derivesFrom(X,wheat).
```

% A1: seafood allergy

```
hasAllergicTrigger(seafood_allergy,X) :- isSubclassOf(X,
seafood).
```

% A2: sesame allergy

```
hasAllergicTrigger(sesame_allergy, sesame_seeds).
```

...

Preparation: substitute proposals

% S1: inferring substitutes from functions

isSubstituteFor(A,B) :- foodProduct(A), foodProduct(B), hasFunction(A,F), hasFunction(B,F).

% S2: inferring substitutes from ML-based recommendations

isSubstituteFor(A,B) :- foodProduct(A), foodProduct(B), ml_similar(A,B,S), threshold(T), S>T.

Model to reason about substitution

%% I: Guess - In the given context, a substitute can be either good or bad

```
goodSubstitute(A,S) | badSubstitute(A,S) :-
isSubstituteFor(A,S).
```

%% II: Check (Constraints)

% C1: exclude substitutes that are allergic triggers

```
:- goodSubstitute(X,S), hasAllergicTrigger(Allergy,
S).
```

%% III: Optimize (Weak constraints)

% W1: if possible, exclude substitutes unacceptable in the given diet(s)

```
:~ badSubstitute(X,S), unacceptableIn(S, Diet). [1]
```

Input/output rules for users

```
% R1: Identifying original ingredients that are wrong

must_replace(X) :- recipe(Recipe), allergy(Allergy),
isNeededFor(X, Recipe), hasAllergicTrigger(Allergy, X).

% R2: Identifying original ingredients that are wrong

should_replace(X) :- recipe(Recipe), diet(Diet),
isNeededFor(X, Recipe), unacceptableIn(X, Diet).

% R3, R4: Projection rules for cleaner output

suggestion(I,S):- must_replace(I), goodSubstitute(I,S).

suggestion(I,S):- should_replace(I), goodSubstitute(I,S).
```

Concluding remarks on reasoning

- ontological modeling and logic programming for knowledge-based substitution
- integration of selected parts of existing ontologies populated with several recipes
- logic program in ASP that allows to reason about recipes and substitutions based on constraints regarding allergies and diets
    - identify wrong ingredients with hard and weak constraints,
    - reason about possible and appropriate substitutes.

Summary and outlook

Research questions revisited

1. Where is the useful knowledge?
   a. domain experts → **elicitation, formalization**
   b. ontologies → **YES! and still room to enrich**
   c. other resources → **datasets, LLM, NER+EL**
2. How to conceptualize/integrate/model it?
   a. **integrate; ontologies + rules + constraints + …**
3. How to reason about it?
   a. **induction (ML)**
   b. **deduction (logic, constraints, *optimization*)**

What we did...

Lots of knowledge engineering tasks:

1. Knowledge acquisition from domain experts
2. Knowledge **integration** reusing existing sources
3. Building **new resources** with NER and EL
4. Knowledge **modeling** and **reasoning**

Open challenges for further research

- Further knowledge base **refinement**, **integration** of new resources, new cases, motivations etc.
- Improving the **knowledge representation** of recipes and substitution (static ingredients + dynamic processes; flavour etc.)
- Expanding the **rule base** about diets, allergies, etc.
- Reasoning about substitution
  - **inferring** substitute candidates
  - constraints and optimization based on **numerical** values

Towards the Internet of Food...

(IC-FOODS)

Acknowledgements and thanks

TAISTI Project: Agnieszka Ławrynowicz (leader, PI), Ania Wróblewska, Jędrzej Potoniec, Jakub Dutkiewicz, Anna Gramza-Michałowska, and more

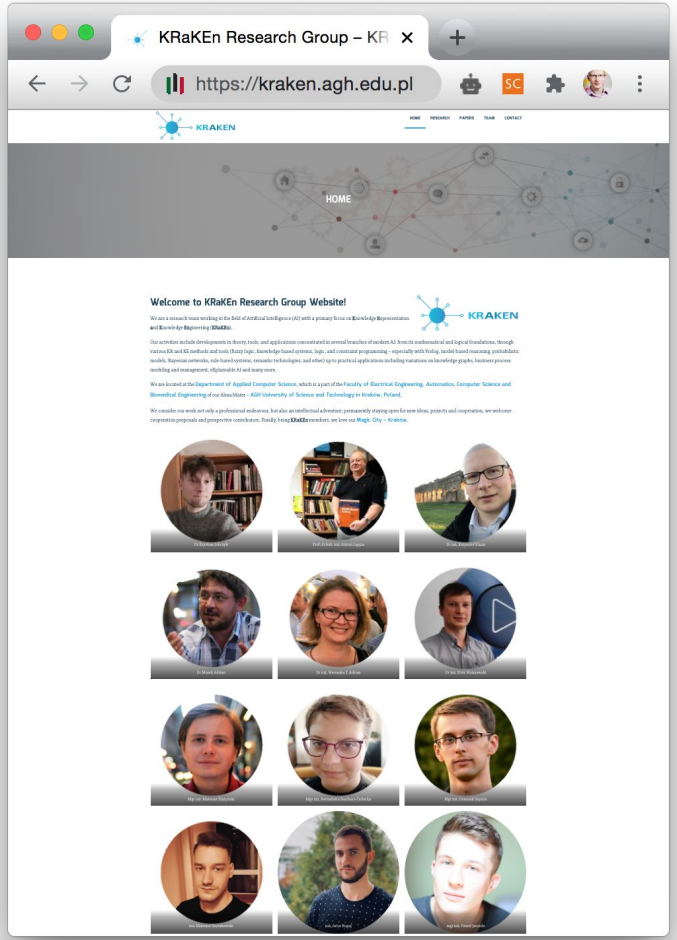
My graduate and undergraduate students (diploma projects/theses):

- Katarzyna Pyrczak, “Żywieniowe grafy wiedzy i wnioskowanie ontologiczne w dziedzinie żywienia”
- Zuzanna Śmiech, “A concept of a substitute in a knowledge graph - review of existing solutions and a practical application in a decision-support system”
- Agata Kawalec, “Combining Natural Language Processing with knowledge representation for finding food substitutes”
- Maciej Kutyła, “Ontologies and Rules for Explainable AI Systems”
- Julia Ignacyk, “Recognition and identification of ingredients using NER and entity linking based on the recipe description”

KRaKEn' Research Group

KRAKEN

@AGH University of Krakow



Thank you for your attention!

Weronika T. Adrian, wta@agh.edu.pl

http://wtadrian.eu, http://kraken.edu.pl