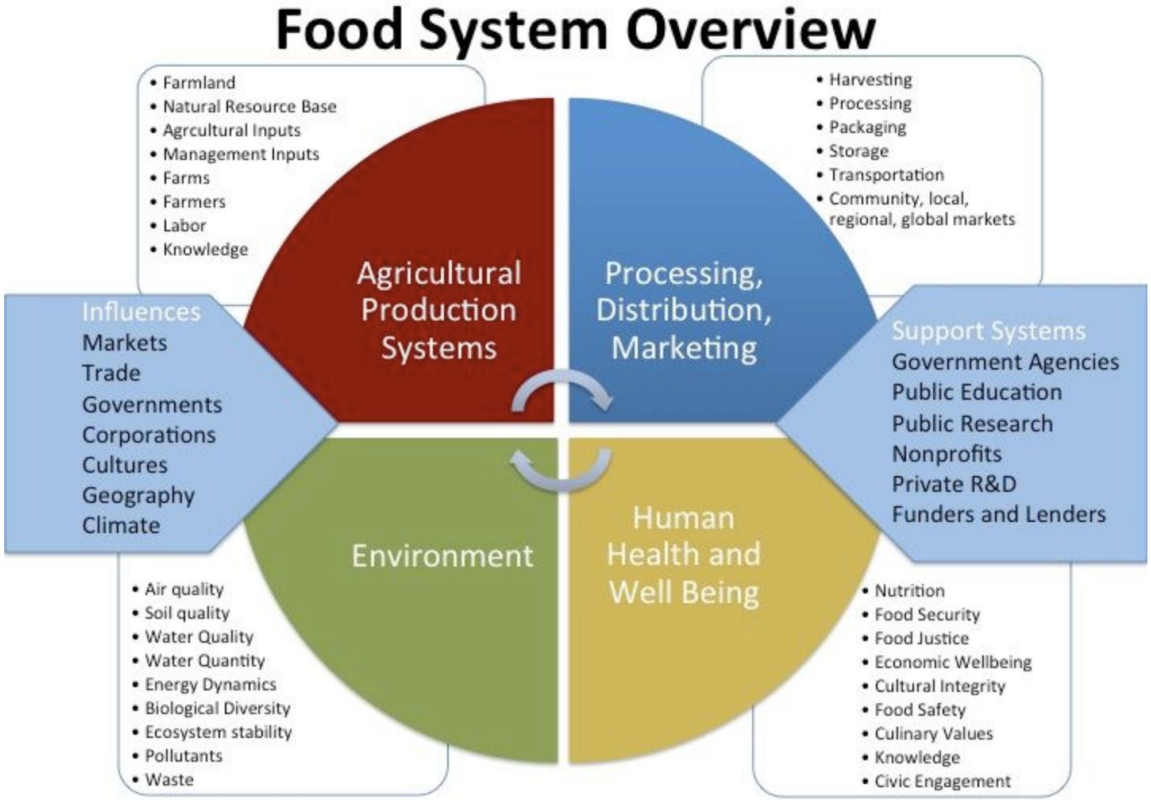
Towards AI-driven Food Science and Society: Opportunities and Challenges

Tome Eftimov and friends

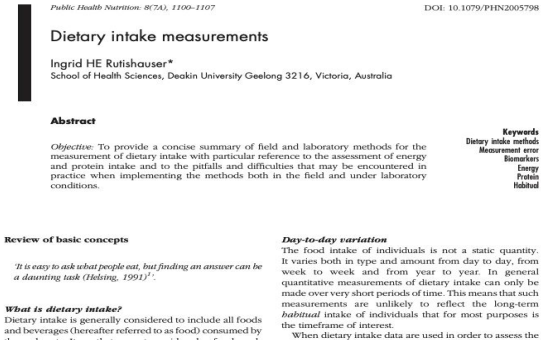
ISO-FOOD - From Food Source to Health
April, 24-26, 2023
Portorož, Slovenia



https://foodsystems.wsu.edu/foodsystemsprogram/

“Big Data” + Digitalization

		Staple Foods							
Nutrients (per 100 gm)	**Unit**	**Maize flour**	**Millet Flour**	**Rice**	**Cassava fresh**	**Cassava flour**	***Matooke* (plantain)**	**Beans**	**Groundnut**
Energy	Kilocalories	369	374	360	160	314	122	347	567
Protein	Grams	7.3	10.9	6.6	1.4	2.6	1.3	21.4	25.8
Fat	Grams	1.8	4.2	0.6	0.3	0.7	0.4	1.2	49.2
Carbohydrate	Grams	79.2	72.1	79.3	38.1	76.6	31.9	62.6	16.1
Calcium	Milligrams	3.0	8.0	9.0	16.0	31.0	3.0	113.0	92.0
Iron	Milligrams	1.1	3.0	0.8	0.3	1.9	0.6	5.1	4.6
Zinc	Milligrams	0.7	1.7	1.2	0.3	0.7	0.1	2.3	3.3



Public Health Nutrition: 8(7A), 1100–1107 DOI: 10.1079/PHN2005798

Dietary intake measurements

Ingrid HE Rutishauser*
School of Health Sciences, Deakin University Geelong 3216, Victoria, Australia

Abstract

Objective: To provide a concise summary of field and laboratory methods for the measurement of dietary intake with particular reference to the assessment of energy and protein intake and to the pitfalls and difficulties that may be encountered in practice when implementing the methods both in the field and under laboratory conditions.

Keywords
Dietary intake methods
Measurement error
Biomarkers
Energy
Protein
Habitual

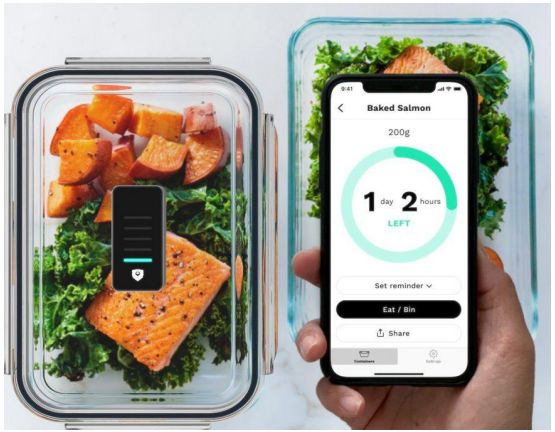
Review of basic concepts

'It is easy to ask what people eat, but finding an answer can be a daunting task (Helsing, 1991)[1].'

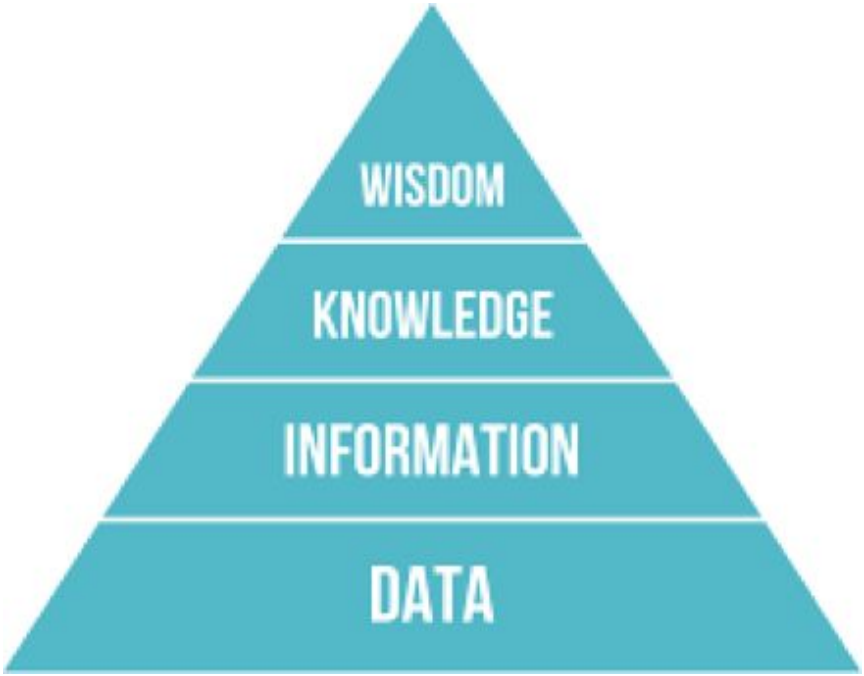
What is dietary intake?
Dietary intake is generally considered to include all foods and beverages (hereafter referred to as food) consumed by

Day-to-day variation
The food intake of individuals is not a static quantity. It varies both in type and amount from day to day, from week to week and from year to year. In general quantitative measurements of dietary intake can only be made over very short periods of time. This means that such measurements are unlikely to reflect the long-term *habitual* intake of individuals that for most purposes is the timeframe of interest.

When dietary intake data are used in order to assess the



Data, Information, Knowledge



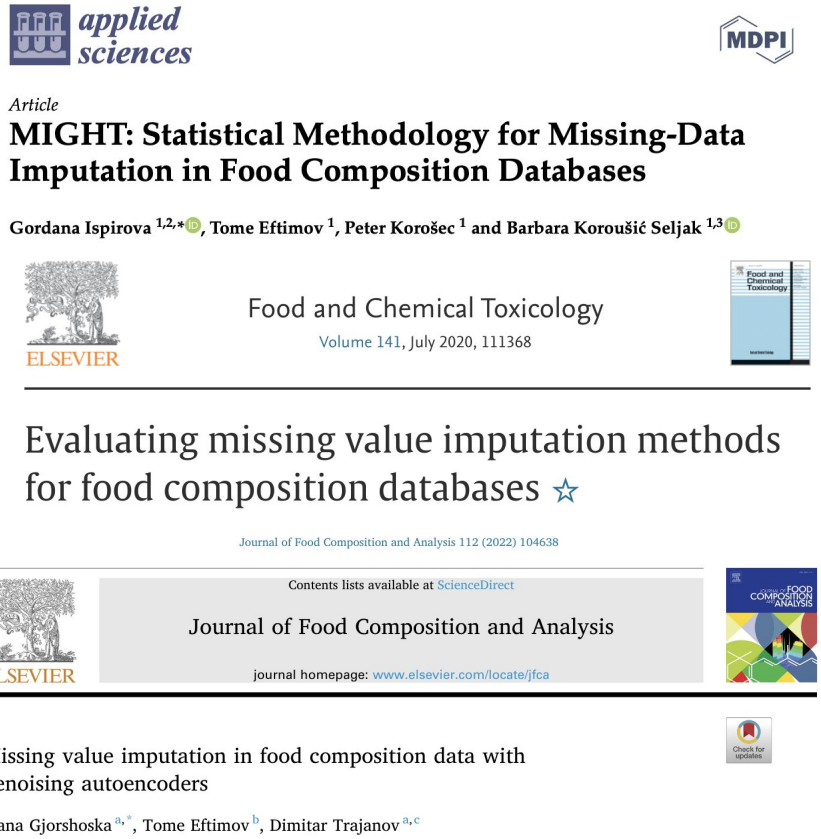
Outline

- **Data**
  - Missing value imputation in Food Composition Databases
- **Information**
  - Extracting food information from scientific literature
- **Knowledge**
  - Food, Chemical, Disease Knowledge Graph Construction
  - Food authenticity
  - Covid-19 mortality rate prediction

Data

Missing value imputation in FCDB

- **Classical approach**
  - Mean or median value of the values of the given nutrient in the same food from several other FCDBs
- **Imputation of a single missing nutrient value**
  - MIGHT - discover from which countries we can borrow using statistical analyses
  - Data mining methods
- **Imputation of multiple missing nutrients values**
  - More realistic scenario - denoising autoencoders



applied sciences

MDPI

Article

MIGHT: Statistical Methodology for Missing-Data Imputation in Food Composition Databases

Gordana Ispirova [1,2,*], Tome Eftimov [1], Peter Korošec [1] and Barbara Koroušić Seljak [1,3]

ELSEVIER

Food and Chemical Toxicology

Volume 141, July 2020, 111368

Evaluating missing value imputation methods for food composition databases ☆

Journal of Food Composition and Analysis 112 (2022) 104638

ELSEVIER

Contents lists available at ScienceDirect

Journal of Food Composition and Analysis

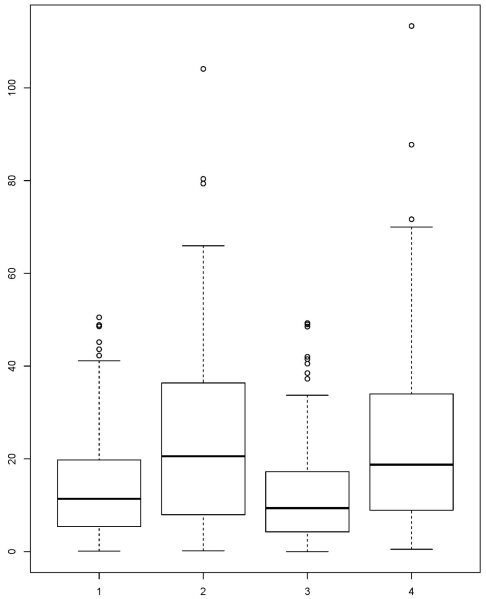
journal homepage: www.elsevier.com/locate/jfca

Missing value imputation in food composition data with denoising autoencoders

Ivana Gjorshoska [a,*], Tome Eftimov [b], Dimitar Trajanov [a,c]

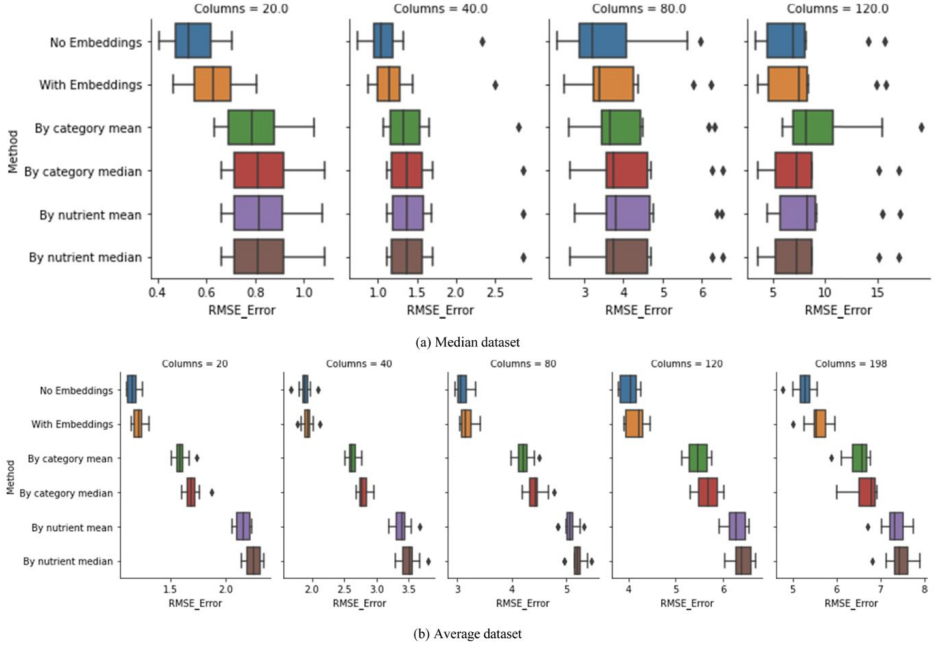
Missing value imputation

Might results



Distribution of absolute error of: MIGHTv1 average, regular average, MIGHTv1 median and regular median calculated for the Potassium (K) content in foods from the food group “Fruits”.

Autoencoders results



Error for 10 iterations (20% missingness)

Information

Extracting food information from scientific literature

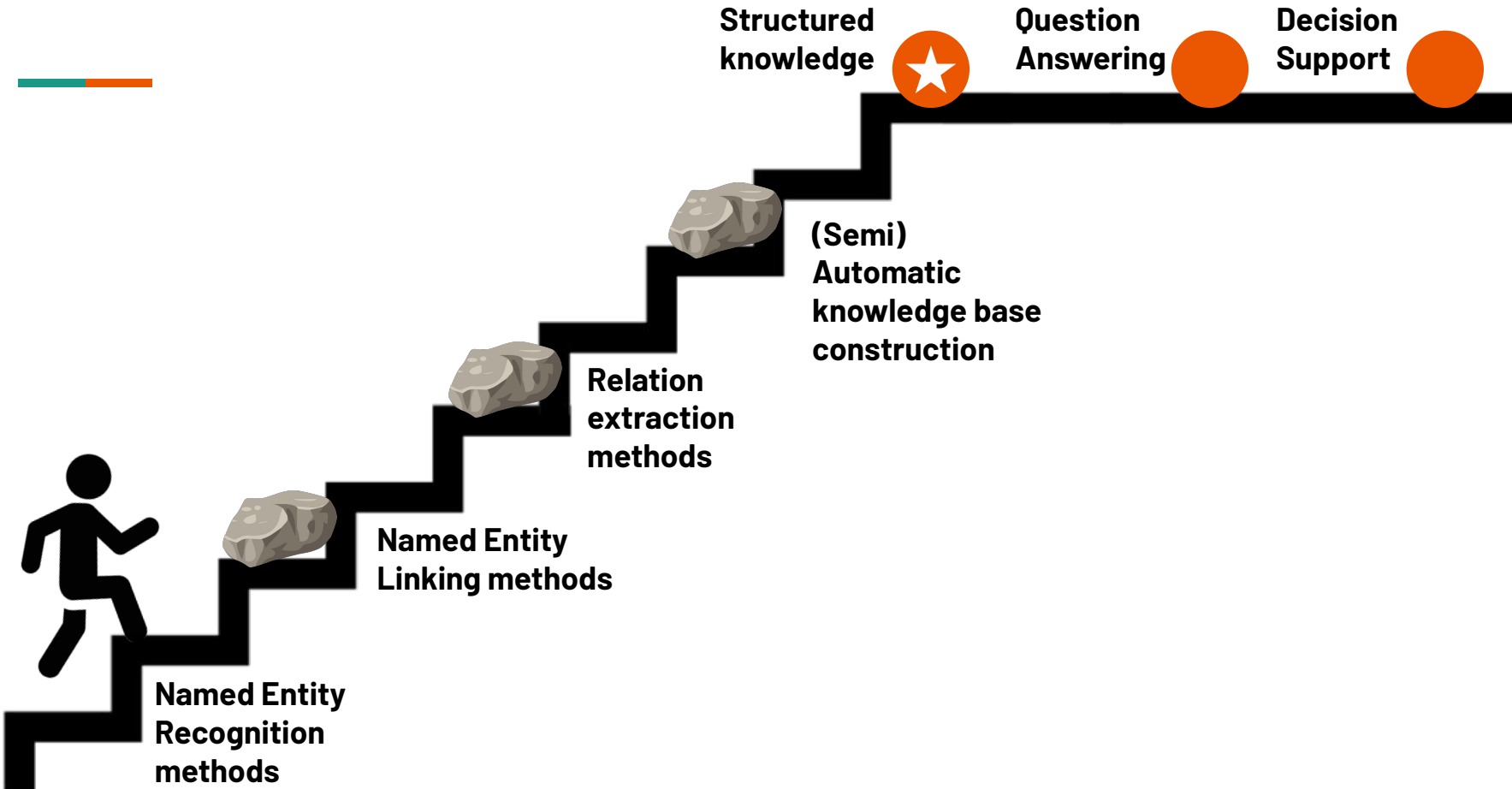
Named-Entity Recognition (NER)

Excessive salt intake has been associated with a higher incidence of heart disease.

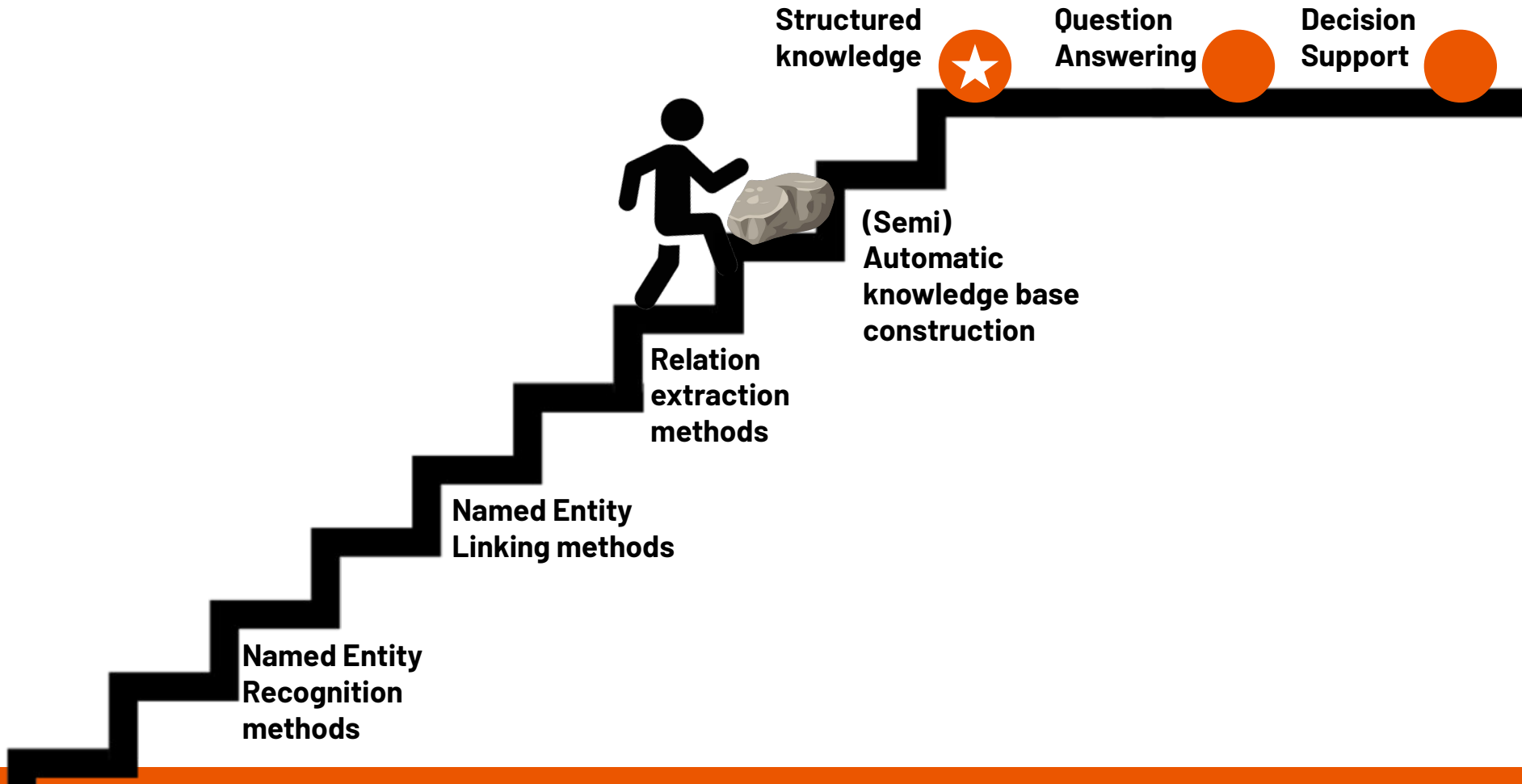
Named-Entity Linking (NEL)

Excessive salt [00002 (FOODB)] intake has been associated with a higher incidence of heart disease [0001 (UMLS)].

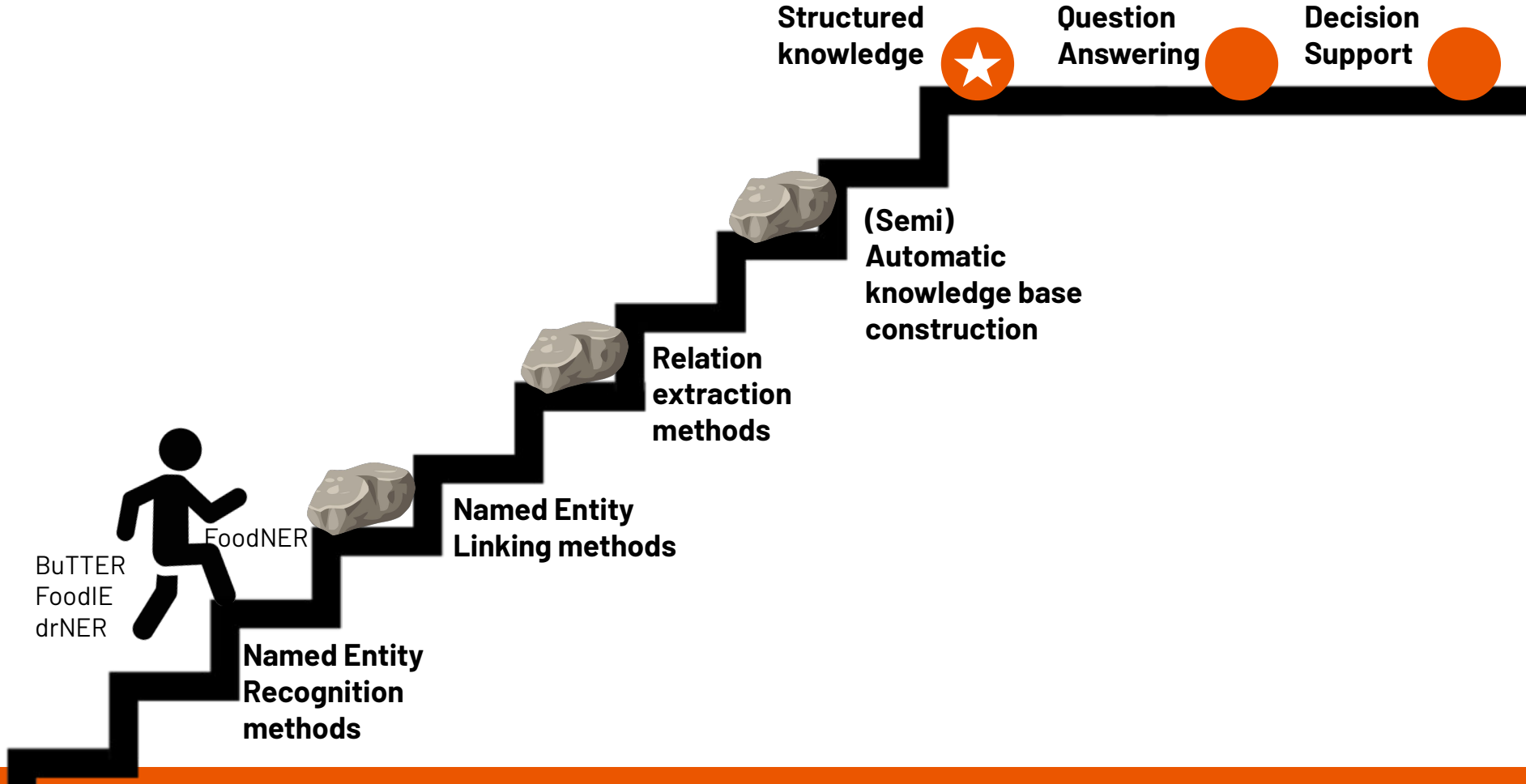
From Language Technologies to Decision Support



State in the Biomedical Domain



State in the Food Domain



FoodViz Tool

Recipes

Currated? ☑

Filter recipes

All categories

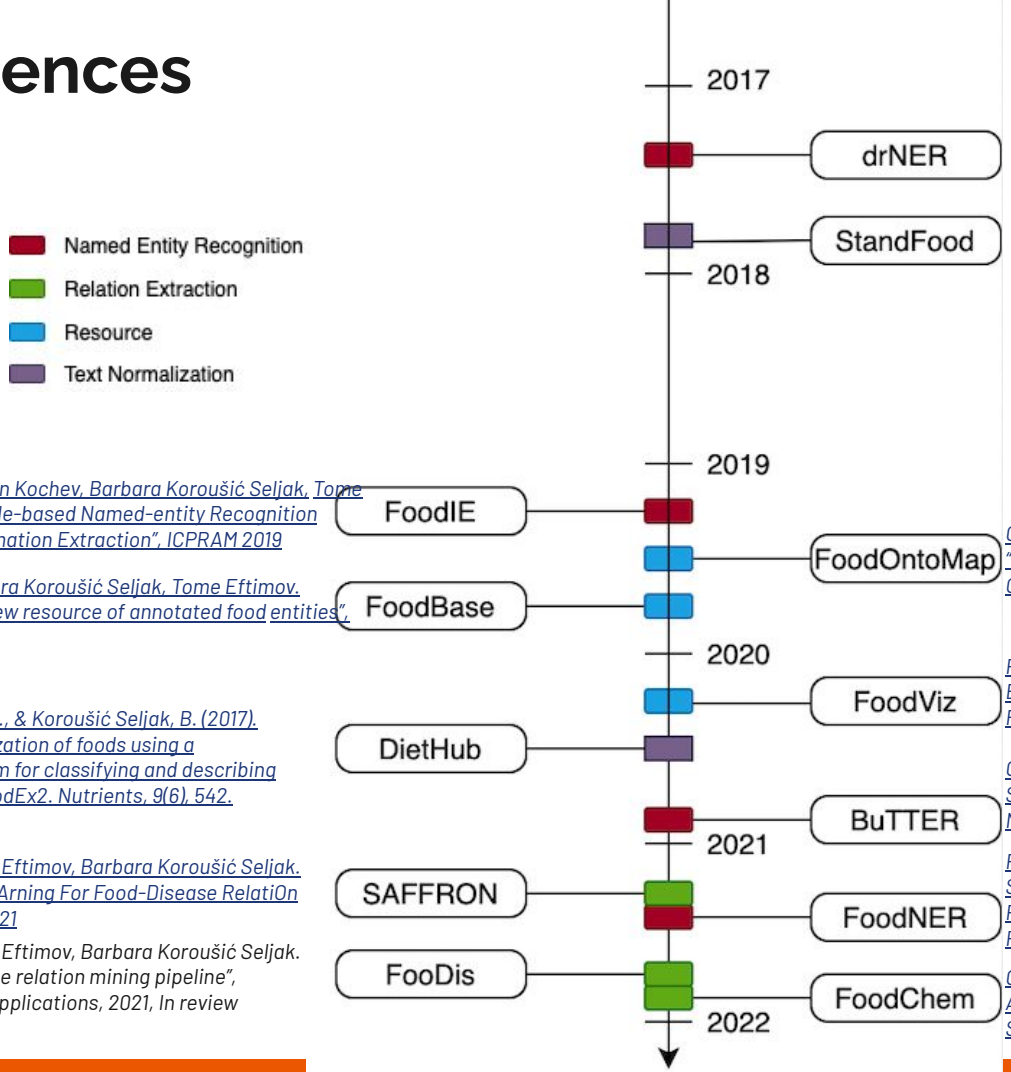
Recognized Entities for recipe **0recipe1006**

Mix the cream cheese, beef, olives, onion, and Worcestershire sauce together in a bowl until evenly blended . Keeping the mixture in the bowl , scrape it into a semi-ball shape . Cover , and refrigerate until firm , at least 2 hours . Place a large sheet of waxed paper on a flat surface . Sprinkle with walnuts. Roll the cheese ball in the walnuts until completely covered . Transfer the cheese ball to a serving plate , or rewrap with waxed paper and refrigerate until needed .

Entity tags

Entity	Synonyms	Hansard Tags	Hansard Parent	Hansard Closest	FoodOn	SnomedCT	OF
cream cheese	CREAM CHEESE	AG.01.e [Dairy produce];AG.01.e.02 [Cheese];AG.01.n [Dishes and prepared food];AG.01.n.18 [Preserve];	Dishes and prepared food	Dairy produce	cream cheese	Cream cheese Cheese Cream	
beef	BEEF	AG.01.d.03 [Beef];	Animals for food	Food		Beef	
olives	OLIVES	AG.01.h.01.e [Fruit containing stone];	Fruit and vegetables	Fruit containing stone		Olives	
onion	ONION	AG.01.h.02.e [Onion/leek/garlic];	Fruit and vegetables	Onion/leek/garlic	onion (whole) Allium cepa	Onion	of:Onion

References



Tome Eftimov, Barbara Koroušić Seljak, Peter Korošec. "A rule-based named-entity recognition method for knowledge extraction of evidence-based dietary recommendations", PLoS One, 2017

Eftimov, T., Korošec, P., & Koroušić Seljak, B. (2017). StandFood: standardization of foods using a semi-automatic system for classifying and describing foods according to FoodEx2. Nutrients, 9(6), 542.

Gorjan Popovski, Stefan Kochev, Barbara Koroušić Seljak, Tome Eftimov. "FoodIE: A Rule-based Named-entity Recognition Method for Food Information Extraction", ICPRAM 2019

Gorjan Popovski, Barbara Koroušić Seljak, Tome Eftimov. "FoodOntoMap: Linking Food Concepts across Different Food Ontologies", KEOD, 2019

Gorjan Popovski, Barbara Koroušić Seljak, Tome Eftimov. "FoodBase corpus: a new resource of annotated food entities", Database, 2019

Riste Stojanov, Gorjan Popovski, Nasi Jofce, Dimitar Trajanov, Barbara Koroušić Seljak, Tome Eftimov. FoodViz: Visualization of Food Entities Linked Across Different Standards, LNCS, 2020

Eftimov, T., Korošec, P., & Koroušić Seljak, B. (2017). StandFood: standardization of foods using a semi-automatic system for classifying and describing foods according to FoodEx2. Nutrients, 9(6), 542.

Gjorgjina Cenikj, Gorjan Popovski, Riste Stojanov, Barbara Koroušić Seljak, Tome Eftimov. "BuTTER: BidirecTional LSTM for Food Named-Entity Recognition", BFNDMA 2020

Gjorgjina Cenikj, Tome Eftimov, Barbara Koroušić Seljak. "SAFFRON: tranSfer leArning For Food-Disease RelatiOn extractioN", NAACL 2021

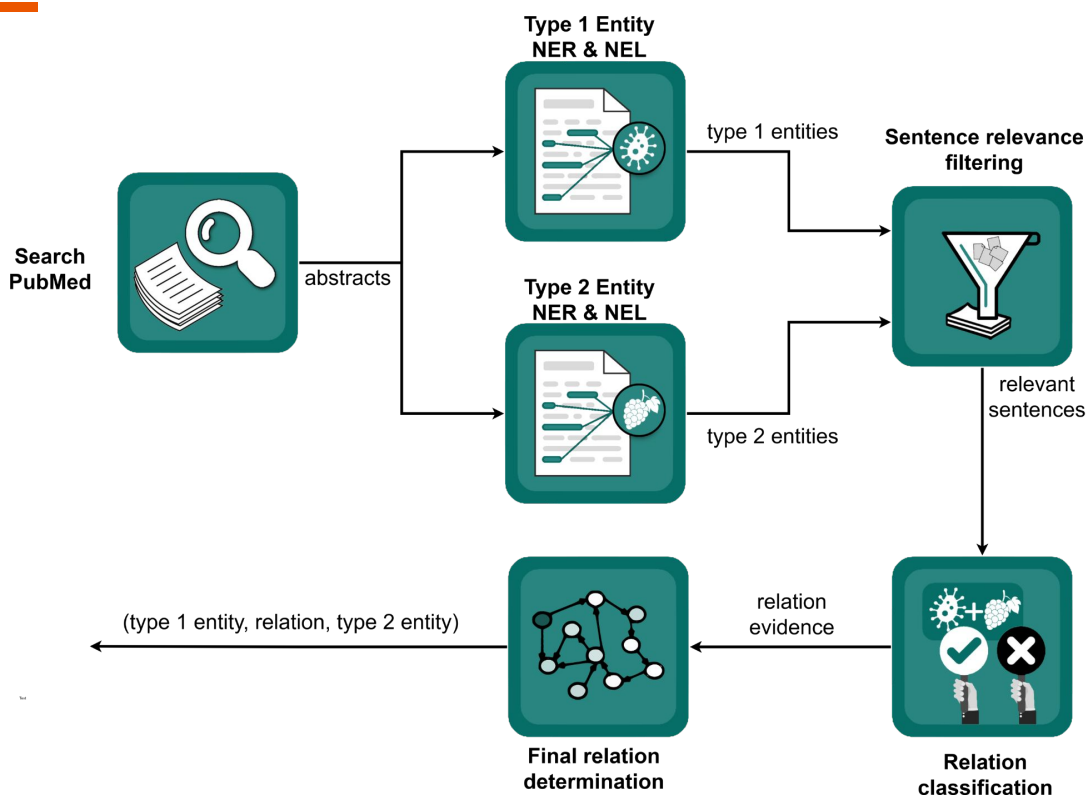
Riste Stojanov, Gorjan Popovski, Gjorgjina Cenikj, Barbara Koroušić Seljak, Tome Eftimov. "A Fine-Tuned Bidirectional Encoder Representations From Transformers Model for Food Named-Entity Recognition: Algorithm Development and Validation", JMIR 2021

Gjorgjina Cenikj, Tome Eftimov, Barbara Koroušić Seljak. "FooDis: A food-disease relation mining pipeline", Expert Systems with Applications, 2021, In review

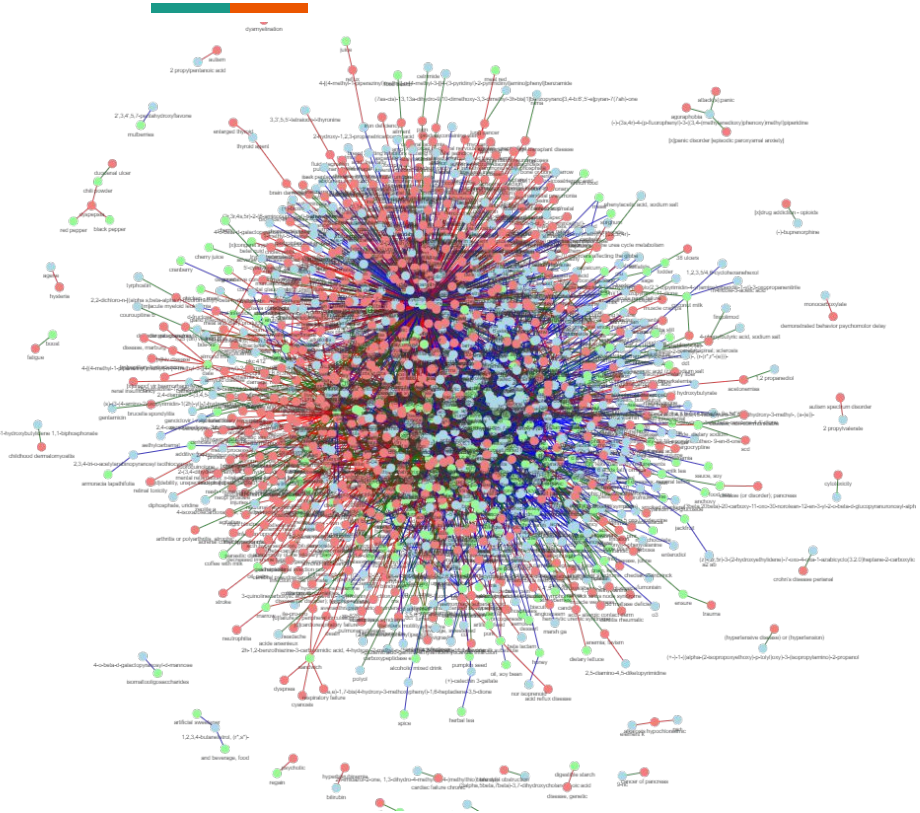
Cenikj, G., Seljak, B. K., & Eftimov, T. (2021, December). FoodChem: A food-chemical relation extraction model. In 2021 IEEE Symposium Series on Computational Intelligence (SSCI)(pp. 1-8). IEEE.

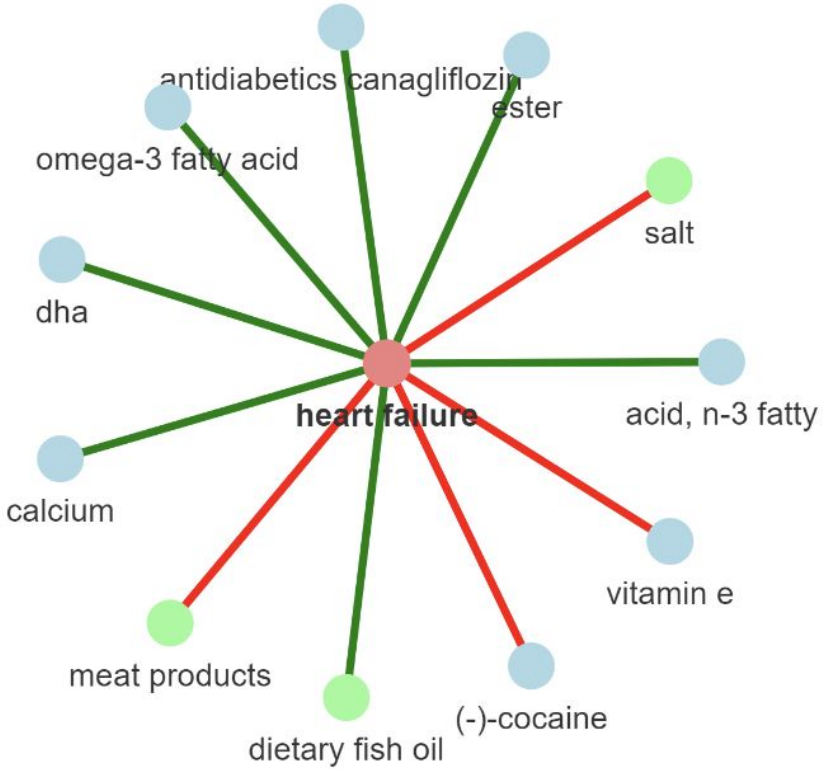
Knowledge

Food, Chemical, Disease Knowledge Graph

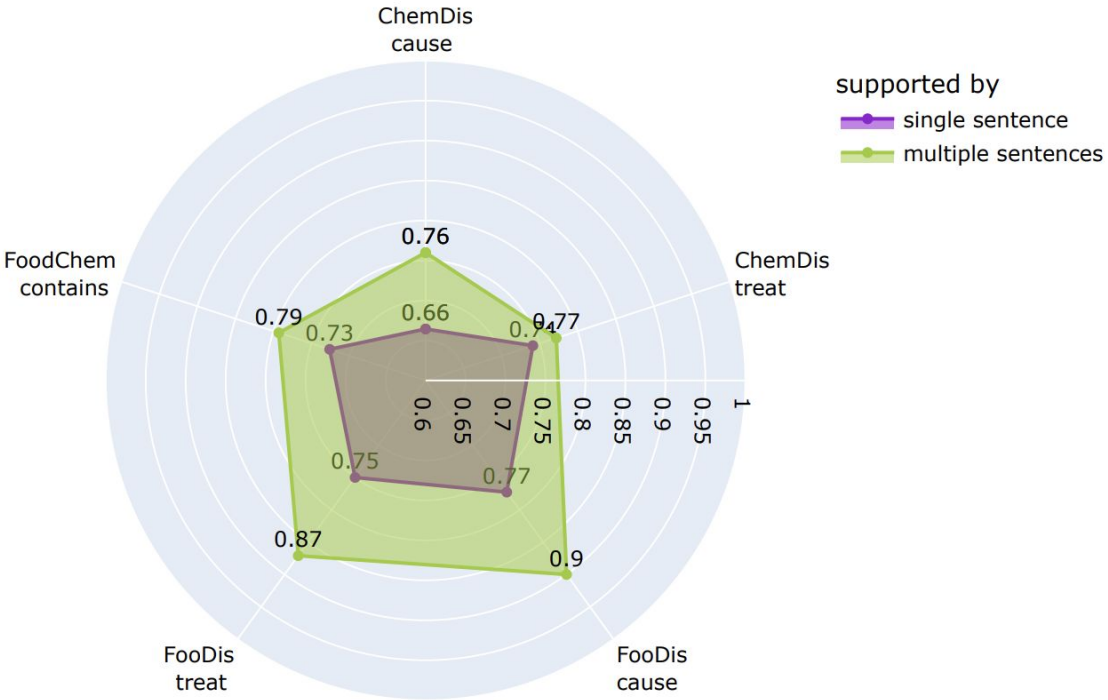


Food, Chemical, Disease Knowledge Graph



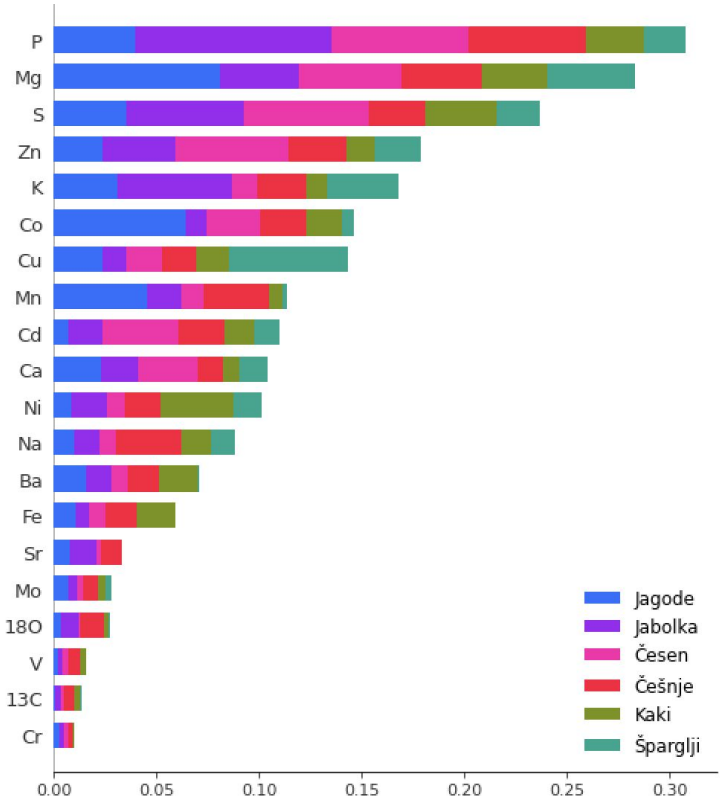


Food, Chemical, Disease Knowledge Graph



Food authenticity

- Based on isotope composition predict the fruit type
- Do we need to use all analyzed experimental sample?
  - Select representative subsample and provide robust results
- Black-box AI models?
- Explainable results -> Providing trust to users

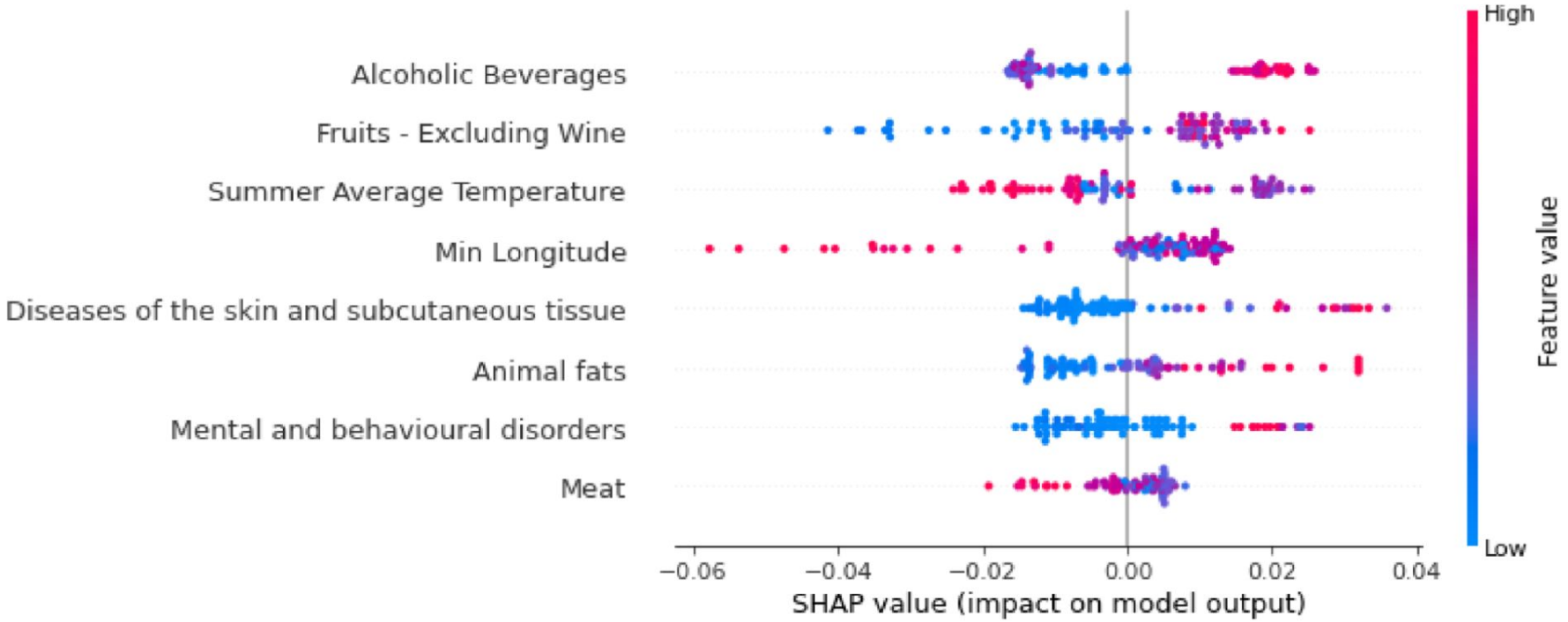


COVID-19 mortality rate

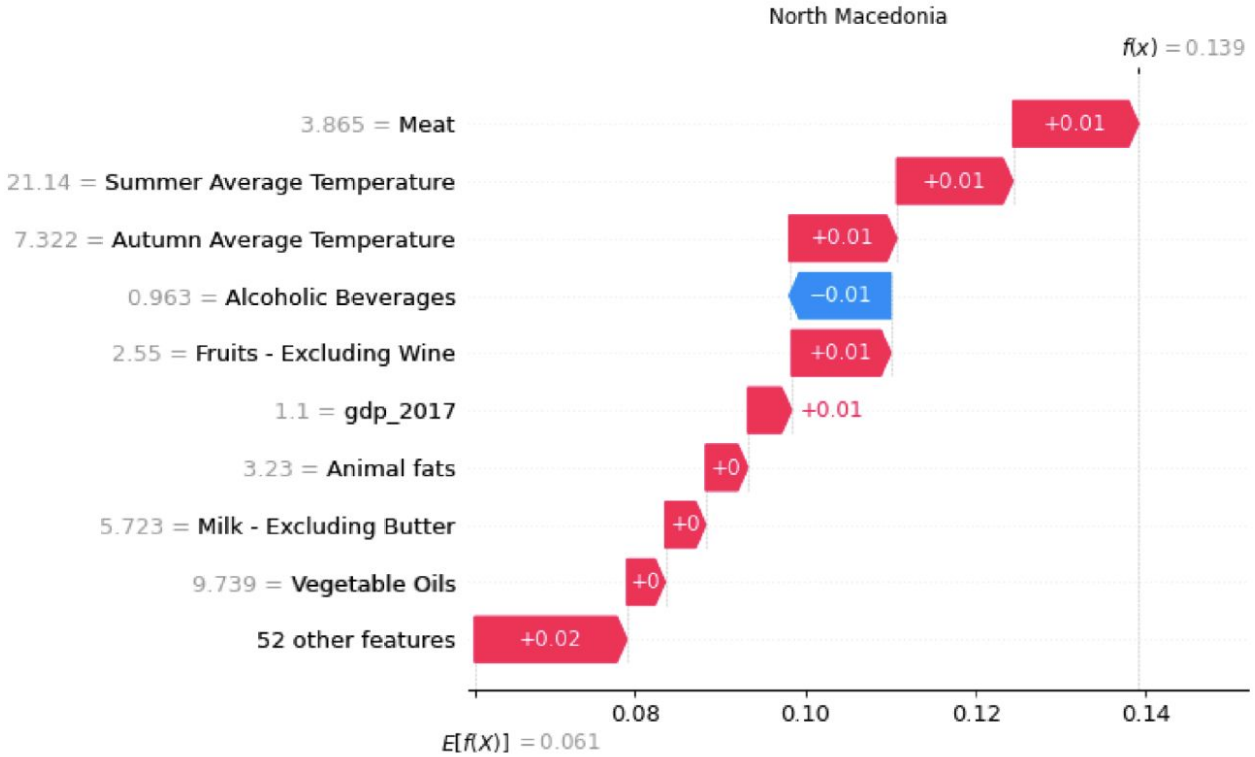
- Predict covid-19 mortality rate on a country level based on the dietary habits, the most common comorbidities, and the socio-economic factors of the country
  - Food consumption - FAO
  - The most common comorbidities - WHO
  - Socio-economic factors
    - Longitude
    - Latitude
    - Average temperature per season
    - GDP
    - ...

Trajanoska, M., Trajanov, R., & Eftimov, T. (2022). Dietary, comorbidity, and geo-economic data fusion for explainable COVID-19 mortality prediction. *Expert Systems with Applications*, *209*, 118377.

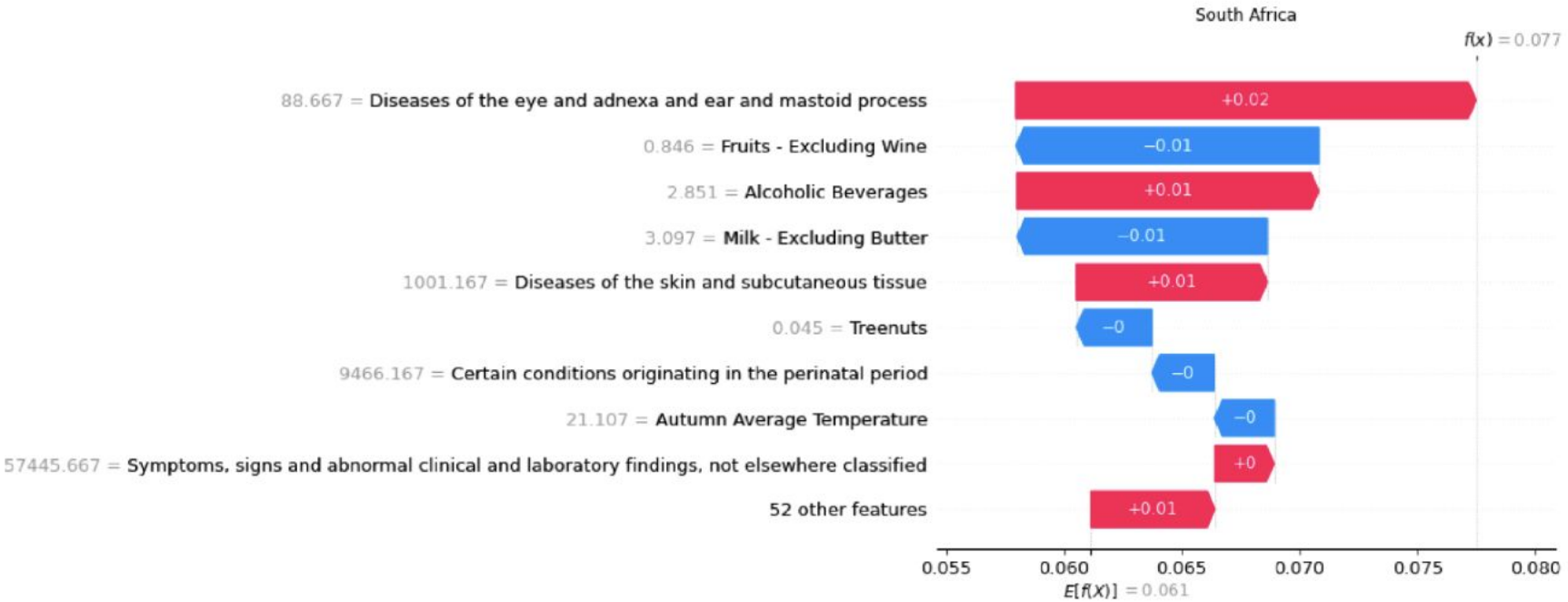
COVID-19 mortality rate



COVID-19 mortality rate



COVID-19 mortality rate



COVID-19 mortality rate

Slovenia

Team

Gjorgjina Cenikj
PhD Student
JSI

Gordana Ispirova, PhD
Postdoctoral researcher
JSI

Matevž Ogrinc
Master Student
JSI

Ana Nikolikj
Master Student
JSI

Eva Valenčič
PhD Candidate
JSI

Asst. Prof. Riste Stojanov, PhD
FINKI

Prof. Peter Korošec, PhD
Senior Researcher
JSI

Asst. Prof. Tome Eftimov, PhD
Senior Researcher
JSI

Prof. Barbara Koroušić Seljak, PhD
Senior Researcher
JSI

Thanks to ...

Public Scholarship, Development, Disability and Maintenance Fund of the Republic of Slovenia