CREATING EUROPEAN FOODOME KNOWLEDGE GRAPH TO CAPTURE THE DARK MATTER OF NUTRITION

Sebek M[1], Jánosov M[2], Ruppert P[1], Józwiak-Hagymásy J[3], Széles G[3], Mészáros Á[3], Országh E[3], Zentai A[3], Gyimesi L[3], Jozwiak A[3], Barabási L[1]

[1] Network Science Institute and Department of Physics, Northeastern University, Boston MA, USA

[2] Geospatial Data Consulting, Budapest, Hungary

[3] Syreon Research Institute, Budapest, Hungary

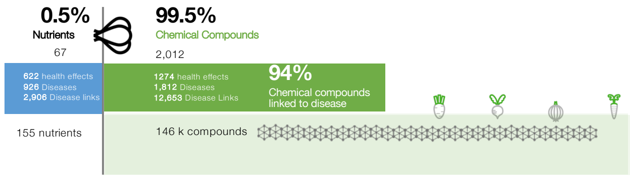
Corresponding author: akos.jozwiak@syreon.eu

MSR185

INTRODUCTION

- A significant part of global public health problems can be traced back to environmental and lifestyle factors.
- Our current understanding of the way food molecules affect health is largely limited to a few hundred components tracked by food composition databases.
- Yet, **these components represent only a tiny fraction of the total composition of the food supply**, hindering research from discovering mechanistic effects and roles of food molecules.
- **To solve this problem, we have designed a Big Data and AI strategy for the creation of a high-resolution collection of food composition: Foodome (from food exposome)**

The Dark Matter of Nutrition



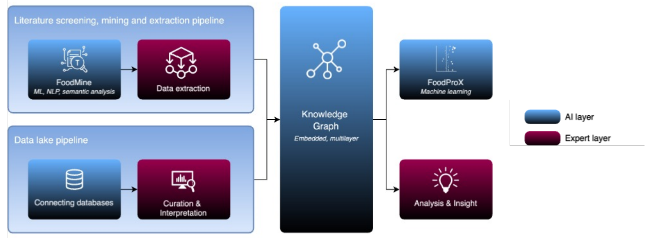
OBJECTIVES

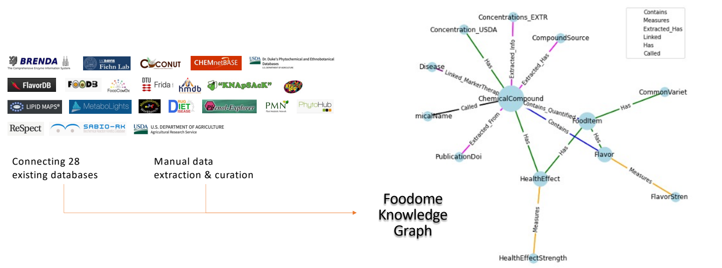
- US Foodome: technology, Nutritional Dark Matter (NDM) database
- AI revolution
- Bringing Foodome to the age of AI while extending to Europe
- Objective: **build a Knowledge Graph** and feed it with data
  - Find and rank relevant information and resources
  - Find relevant terms
  - Connect existing dataset
  - Develop a data lake where entities are connected

METHODS

- In this European pilot conducted in Hungary, data from **28 existing** food composition, biochemical, and health effect **databases were connected** and curated.
- An **AI-assisted pipeline (FoodMine)** was designed to complement the database curation, reviewing and extracting data from scientific literature for the 8 most relevant Hungarian food commodities.
- The resulting database was designed to **show the interplay between foods**, their **compounds**, and **potential downstream effects** of their consumption through a network science approach.

Foodome layers

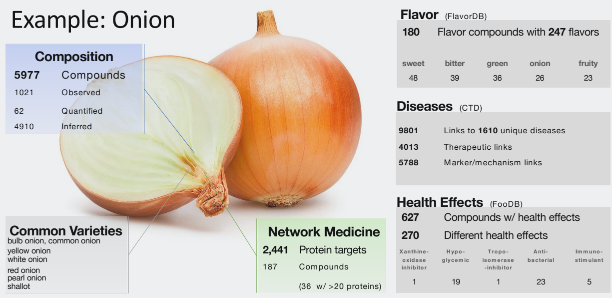




RESULTS

- A unified Hungarian **Foodome Knowledge Graph (FKG)** was created.
- The FKG contains ontologies of 10 different entity types (FoodItem, CommonVariety, ChemicalCompound, ChemicalName, Concentration, Flavor, FlavorStrength, Disease, HealthEffect, HealthEffectStrength) and 20 relationships (e.g., chemical-disease associations, food item-health effect associations).
- This pilot FKG alone contains **30k+ entities** and **100k+ relationships**, revealing connections between food compounds and potential mechanisms of action, both marker and therapeutic, and disease associations as well.

Example: Onion



IMPACT

Nutrition science: demonstrating that today's nutritional recommendations are not (or only partially) evidence-based
Example: Integrating recent advances of medical sciences

FoodAsDrug: Identify food molecules with beneficial health effects
Example: Mapping the health implications of 53 polyphenols

Food regulatory framework: developing new public health regulatory legislation
Example: Incentive system based on composition or health effects

Start-up ecosystems: agri-food, nutrition, health industry can all benefit from results
Example: Personalised evidence-based dietary app

Health science: developing nutritional guidelines and individual-targeted (personalized) nutrition applications
Example: Developing an app showing the processing level of foods

Food&Drug Intractions: Predict foods that alter the efficacy of specific drugs
Example: Why avocado decreases the effectiveness of Warfarin

Agri-food industry: enabling the development of functional/healthy/wholesome food products
Example: Development of functional foods based on health evidence

AI: demonstrating the applicability of cutting-edge AI technology in evidence generation
Example: AI-based relevance screening and data extraction

REFERENCES

Barabási, AL., Menichetti, G. & Loscalzo, J. The unmapped chemical complexity of our diet. Nat Food 1, 33–37 (2020).
Milanlouei, S., Menichetti, G., Li, Y. et al. A systematic comprehensive longitudinal evaluation of dietary factors associated with acute myocardial infarction and fatal coronary heart disease. Nature Communications, 11(1), 6074. (2020).
Hooton, F., Menichetti, G. & Barabási, AL. Exploring food contents in scientific literature with FoodMine. Sci Rep 10, 16191 (2020).
do Valle, I.F., Roweth, H.G., Malloy, M.W. et al. Network medicine framework shows that proximity of polyphenol targets and disease proteins predicts therapeutic effects of polyphenols. Nat Food 2, 143–155 (2021).
Menichetti, G. An AI pipeline to investigate the binding properties of poorly annotated molecules. Nature Reviews Physics, 4(6), 359–359 (2022).
do Valle, I. F., Roweth, H. G., Malloy, M. W. et al. Network medicine framework shows that proximity of polyphenol targets and disease proteins predicts therapeutic effects of polyphenols. Nature Food, 2(3), 143–155 (2021).
Menichetti, G., Ravandi, B., Mozaffarian, D. et al. Machine learning prediction of the degree of food processing. Nat Commun 14, 2312 (2023).

John von Neumann University
Funded by the John von Neumann University, Kecskemét, Hungary

Presented at the ISPOR Europe 2024 Conference, 17 – 20 NOVEMBER 2024 BARCELONA, SPAIN

BARABASILAB

SYREON