

An Insight into Food Semantics: Review, Analysis, and Lessons Learnt over Food-related Studies

Gorjan Popovski¹, Gordana Ispirova¹, Eva Valenčič¹,
Riste Stojanov², Tome Eftimov¹, Barbara Koroušič Seljak¹

¹ Computer Systems Department, Jožef Stefan Institute, Jamova cesta 39, 1000 Ljubljana, Slovenia

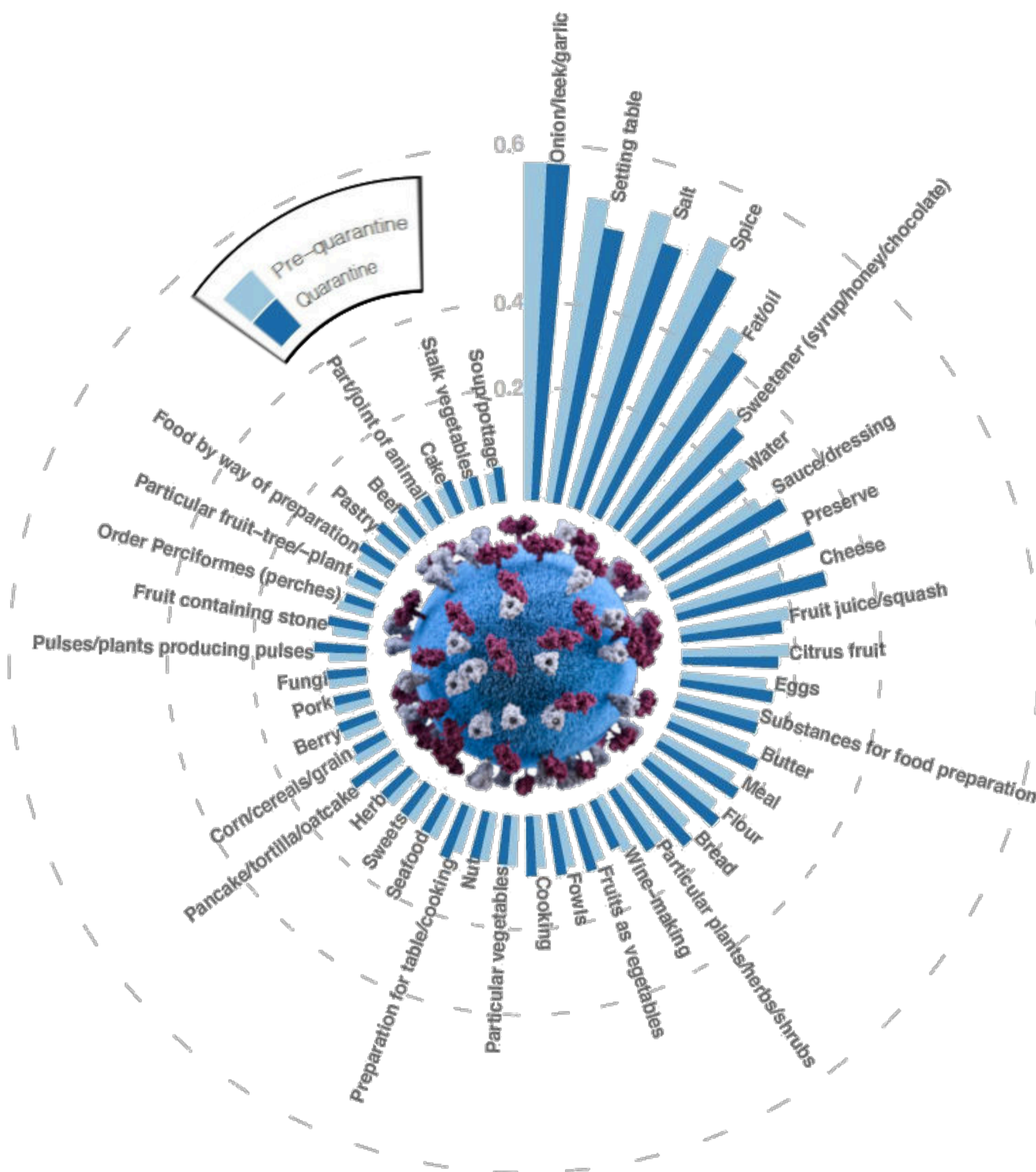
² Faculty of Computer Science and Engineering, Ss. Cyril and Methodius University, Rugier Boshkovikj 16, 1000 Skopje, North Macedonia

Food Information Extraction

- Information Extraction (IE) is the task of automatically extracting structured information from unstructured data (e.g. text).
- Named-Entity Recognition (NER) is a sub-task of IE which focuses on recognizing and extracting entities of interest from unstructured textual data.
- NCBO Annotator using FoodOn, SNOMED CT, and OntoFood.
- Rule-based food NER methods: **DrNER** [1], **FoodIE** [2].
- Study: Comparing FoodIE with NCBO Annotator food NERs. [3]

Food Data Normalization

- **StandFood**: Semi-automatic classification system according to FoodEx2. [4]
- Domain-coverage analysis of a language for describing food. [5]
- **FoodBase**: The first data corpus containing recipes annotated with food entities. [6]
- **FoodOntoMap**: Food concept mapping across different food semantic resources. [7]
- **FoodViz**: A new visualization tool aimed at facilitating subject-matter integration. [8]



Open gaps

- Linking extracted food concepts with concepts from other domains: health, biomedicine, consumer and social sciences, etc.
- Usage to improve personalized nutrition and medicine, as well as public health.

Contact & Funding

- Gorjan Popovski, gorjan.popovski@ijs.si, Jožef Stefan Institute, Jamova cesta 39, 1000 Ljubljana, Slovenia
- This research was supported by the Slovenian Research Agency (research core grant number P2-0098), and the European Union's Horizon 2020 research and innovation programme (FNS-Cloud, Food Nutrition Security) (grant agreement 863059).

References

- 1 T. Eftimov, B. Koroušič Seljak, and P. Korošec, "A rule-based named-entity recognition method for knowledge extraction of evidence-based dietary recommendations," *PLoS One*, vol. 12, no. 6, p. e0179488, 2017.
- 2 G. Popovski, S. Kochev, B. Koroušič Seljak, and T. Eftimov, "Foodie: A rule-based named-entity recognition method for food information extraction," *ICPRAM 2019*, pp. 915–922, 2019.
- 3 G. Popovski, B. Koroušič Seljak, and T. Eftimov, "A survey of named-entity recognition methods for food information extraction," *IEEE Access*, vol. 8, pp. 31586–31594, 2020.
- 4 T. Eftimov, P. Korošec, and B. Koroušič Seljak, "Standfood: Standardization of foods using a semiautomatic system for classifying and describing foods according to FoodEx2," *Nutrients*, vol. 9, no. 6, p. 542, 2017.
- 5 G. Popovski, B. Paudel, T. Eftimov, and B. Koroušič Seljak, "Exploring a standardized language for describing foods using embedding techniques," in *IEEE Big Data 2019*, pp. 5172–5176, 2019.
- 6 G. Popovski, B. Koroušič Seljak, and T. Eftimov, "FoodBase corpus: a new resource of annotated food entities," *Database*, vol. 2019, 11 2019. baz121.
- 7 G. Popovski, B. K. Seljak, and T. Eftimov, "Foodontomap: Linking food concepts across different food ontologies," *IC3K KEOD*, pp. 195–202, INSTICC, SciTePress, 2019.
- 8 R. Stojanov, G. Popovski, N. Jofce, D. Trajanov, B. Koroušič Seljak, and T. Eftimov, "Foodviz: Visualization of food entities linked across different standards," *LOD 2020*, p. In Press., 2022.
- 9 G. Popovski, G. Ispirova, N. Hadzi-Kotarova, E. Valenčič, T. Eftimov, and B. K. Seljak, "Food data integration by using heuristics based on lexical and semantic similarities," *BIOSTEC HEALTHINF 2020*, pp. 208–216, INSTICC, SciTePress, 2020.