Food Informatics: Leveraging data and knowledge to advance food systems

Mabel Zitong Leow,[1] Caili Fu,[2,*] and Dachuan Zhang[1,2,*]
[1]Department of Food Science and Technology, Faculty of Science, National University of Singapore, Singapore 117542, Singapore
[2]National University of Singapore Suzhou Research Institute, Suzhou 215123, China
*Correspondence: caili.fu@nusri.cn (C.F.); dachuan.zhang@nus.edu.sg (D.Z.)
Received: February 4, 2026; Accepted: April 16, 2026; Published Online: April 17, 2026; https://doi.org/10.59717/j.xinn-inform.2026.100038

Citation: Leow M. Z., Fu C. and Zhang D. (2026). Food Informatics: Leveraging data and knowledge to advance food systems. *The Innovation Informatics* **2**:100038.

The term Food Informatics has been increasingly used in scientific literature to describe the growing use of data-driven and computational approaches in food research. However, its definition and scope remain inconsistent across disciplines. This lack of conceptual clarity limits knowledge accumulation and hinders the development of Food Informatics as a coherent interdisciplinary field. In this commentary, we review current definitions of Food Informatics and propose an updated definition with a clearer scope, aiming to provide a conceptual foundation for leveraging data, knowledge, and methods across food systems.

EXISTING DEFINITION AND SCOPE OF FOOD INFORMATICS

The absence of a widely accepted definition has led to a fragmented conceptual landscape in Food Informatics. Early work by Martinez-Mayorga and Medina-Franco[1] framed Food Informatics as an extension of cheminformatics, emphasizing computational tools for managing and analyzing large chemical datasets relevant to food chemistry. As interest broadened, the scope of Food Informatics expanded to encompass more segments of the agri-food system. For example, the Food and Agriculture Organization of the United Nations has advanced agro-informatics to support agricultural digital transformation by integrating agricultural data with information technologies such as geospatial tools and remote sensing, to improve data accessibility and support evidence-based interventions.

More recently, Krupitzer and Stein[2] described Food Informatics as a supporting discipline that applies the Internet of Things and artificial intelligence (AI) across the food production chain, including agriculture, logistics, processing, retail, consumption, and waste management, while incorporating elements of circular-economy thinking. This framing represents a key step towards recognizing Food Informatics as a discipline spanning the entire food chain.

Nevertheless, across these evolving usages, Food Informatics has often been treated as a supporting tool for neighboring areas, such as food safety, supply chain management, or agriculture, rather than being developed as a discipline with its own core research questions and methods. While these definitions emphasize different aspects of the field, they still leave room for further clarification regarding the conceptual focus and scope of Food Informatics. As a result, research has often progressed in parallel at different scales, with relatively limited cross-scale integration of data and methods.

Importantly, differences among existing definitions largely reflect variations in analytical perspective rather than fundamental conceptual disagreement. Current usages of Food Informatics emphasize different levels of analysis, including molecular-level information (e.g., chemical composition and activity), food-level properties (e.g., formulation, processing, and sensory attributes), or system-level aspects such as supply chains and agri-food networks. These perspectives are complementary, as food systems are inherently multi-scale. A definition that privileges only one level risks overlooking important research directions. Instead, Food Informatics can be understood as an integrative framework that connects these different levels of food-related information within a coherent conceptual structure.

UPDATED DEFINITION AND SCOPE OF FOOD INFORMATICS

Building on previous work, we propose the following definition of Food Informatics.

Food Informatics is an interdisciplinary field that acquires, organizes, analyzes, and interprets food-related data and knowledge.

It lies at the intersection of food science, informatics, and computational science, aiming to enable knowledge discovery, decision-making, and optimization across food systems.

Food Informatics operates across multiple interconnected layers of the food system, ranging from molecular and ingredient-level information to food and product-level processes, and further to system-level interactions across the full food lifecycle (Figure 1). A range of computational and data-driven approaches contribute to this framework, including AI for prediction and pattern discovery, the Internet of Things for large-scale data acquisition, and system-level approaches such as life cycle assessment to evaluate environmental impacts. These methods are applied across diverse domains, including food production and security, supply and processing, safety and quality, nutrition and health, perception, consumption and circularity, and beyond.

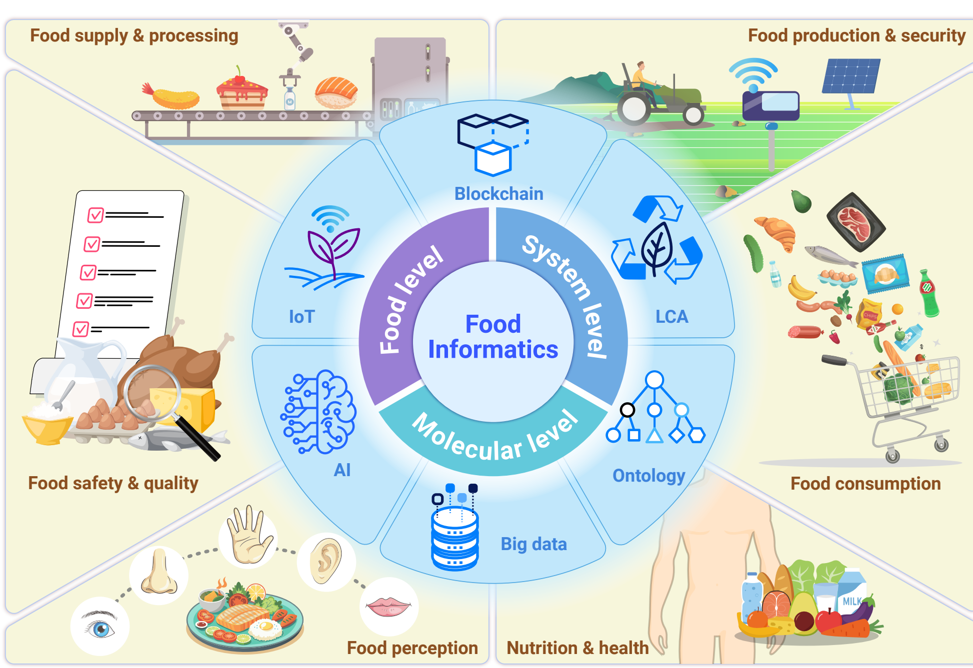

Figure 1. Overview of the scope (central circle), key approaches (intermediate circle), and major research areas (surrounding sectors) in Food Informatics.

CONNECTION AND DIFFERENCES WITH RELATED CONCEPTS

Several concepts have been used to describe computational research related to food, including AI in Food Science, Food Omics, Food Computing, and Computational Gastronomy. These areas share common ground with Food Informatics in their use of data-driven approaches, but differ in their primary scope and focus.

AI in Food Science refers to the application of AI to specific food-related tasks, such as quality assessment, safety evaluation, or process optimization, with an emphasis on predictive performance rather than disciplinary structure. Food Omics focuses on molecular profiling of food systems, such as genomics, proteomics, and metabolomics, to characterize food composition and functionality. Food Computing, originating mainly from computer science and human–computer interaction communities, focuses on task-oriented applications such as food image recognition, dietary monitoring, and food recommendation. Computational Gastronomy examines culinary patterns and structures through data-driven analysis of recipes, flavors, and ingredient combinations, providing insights into gastronomic structure and creativity.

In contrast, Food Informatics is not defined by a particular method, task, or application domain. Instead, it emphasizes the representation, integration, and analysis of heterogeneous food-related information and knowledge across the food system. From this perspective, AI in Food Science, Food Omics, Food Computing, and Computational Gastronomy can be viewed as complementary research areas that address specific application domains, while Food Informatics provides a broader conceptual framework for connecting and leveraging these diverse forms of food-related information.

DATA-DRIVEN INSIGHTS ENABLED BY FOOD INFORMATICS

Representative studies across molecular, food, and system scales illustrate the types of insights enabled by Food Informatics. At the molecular level, the integration of large food compound databases has enabled systematic exploration of food chemical space. Analysis of more than 70,000 food-related molecules showed that food chemicals occupy a broader and more heterogeneous chemical space than common natural products or approved drugs alone.[1] These findings highlight the remarkable chemical diversity of food and the value of informatics approaches for understanding food composition at scale.

At the food and product level, integrating image-based and instrumental data has enabled more objective characterization of food flavor, functions, and quality. Machine-learning models applied to food images or omics data have achieved high accuracy in classifying and grading fruits such as tomatoes, apples, and strawberries, demonstrating the potential of data-driven approaches to enable scalable, reproducible assessment beyond manual inspection.[3]

At the system level, large-scale integration of life cycle assessment datasets has revealed substantial variability in environmental impacts across the global food supply chain. Analysis of approximately 38,700 farms showed up to 50-fold differences in greenhouse gas emissions among producers of the same product, with the highest-impact 25% responsible for more than half of total emissions for many foods.[4] Together, these examples illustrate how Food Informatics can generate new insights by integrating heterogeneous data, contributing to a more comprehensive understanding of food systems.

FUTURE DIRECTIONS

As Food Informatics continues to mature, several priorities will be important for its development as a coherent interdisciplinary field. While existing studies have generated valuable insights, many still focus on a single level of analysis, such as molecular composition, food properties, or system-level analysis. Future progress will therefore benefit from integrative approaches that connect these levels and enable coordinated analysis of molecular, food, and system-level information. Such cross-scale integration would make it possible to address questions that cannot be resolved within a single level of analysis. For example, linking food chemical composition to sensory perception and nutritional outcomes could enable the prediction of flavor, texture, or health effects directly from molecular composition, thereby supporting data-driven food design. Similarly, connecting product-level characteristics with system-level datasets could reveal how specific food formulations or processing choices translate into environmental impacts, such as greenhouse gas emissions or resource use, across supply chains, thereby informing the development of more effective sustainability interventions. Developing robust approaches for knowledge representation and data fusion, including ontologies, knowledge graphs, and hybrid data–knowledge models, will be particularly important for supporting cross-scale analysis.

Strengthening data foundations is another key priority. Many existing datasets or models remain fragmented, heterogeneous, and inconsistently curated, which can limit their reuse and integration. Advancing Food Informatics will require systematic data curation, transparent data provenance, and interoperable standards that enable molecular, food-level, and system-level data to be linked and analyzed together.[5]

In addition, future research may increasingly shift from prediction-oriented applications toward discovery- and design-oriented workflows. Beyond building new models and improving model accuracy, Food Informatics has the potential to support hypothesis generation, exploration of food design spaces we cannot reach with conventional trial-and-error approaches, and evaluation of trade-offs among quality, nutrition, safety, and sustainability.

Closer integration between computational modeling, experimentation, and domain expertise will also be important. Human–AI collaboration and human-in-the-loop workflows that connect modeling with experimentation and product development will play an increasingly important role in the food industry, where data collection is costly and expert knowledge remains essential.

Food Informatics can further contribute to decision support and governance by translating complex, multi-scale data into actionable insights for policymakers, industry stakeholders, and consumers.

Finally, the long-term development of Food Informatics depends on education and training. Dedicated courses or degree programs have begun to emerge at institutions such as the National University of Singapore, Wageningen University & Research, and the University of Hohenheim. Expanding such programs to additional institutions will help make training in Food Informatics more accessible and support a new generation of researchers capable of advancing Food Informatics as a distinct interdisciplinary field.

REFERENCES

1. Martinez-Mayorga K. and Medina-Franco J.L. (2014). Foodinformatics: Applications of chemical information to food chemistry. Martinez-Mayorga K. and Medina-Franco J.L. (eds). (Springer). DOI:10.1007/978-3-319-10226-9
2. Krupitzer C. and Stein A. (2021). Food Informatics — Review of the current state-of-the-art, revised definition, and classification into the research landscape. *Foods* **10**:2889. DOI:10.3390/foods10112889
3. Ke Q., Zhang J., Huang X., et al. (2026). Machine learning unveils three layers of food complexity. *npj Sci. Food* **10**:87. DOI:10.1038/s41538-026-00730-w
4. Poore J. and Nemecek T. (2018). Reducing food's environmental impacts through producers and consumers. *Science* **360**:987–992. DOI:10.1126/science.aaq0216
5. Zhang D., Liu M., Yu Z., et al. (2025). Domain knowledge, just evaluation, and robust data standards are required to advance AI in food science. *Trends Food Sci Technol* **164**:105272. DOI:10.1016/j.tifs.2025.105272

FUNDING AND ACKNOWLEDGMENTS

This research was supported by the Ministry of Education, Singapore, under the Academic Research Fund Tier 1 (A-8003718-00-00) and the Start-Up Grant of the National University of Singapore (A-0010237-00-00). The authors used AI tools for language polishing and grammar checks and are fully responsible for the content and conclusions. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

D.Z. and M.Z.L. designed the research. M.Z.L. and D.Z. wrote the paper with input from C.F. All authors contributed to the manuscript and approved the final version.

DECLARATION OF INTERESTS

The authors declare no competing interests.

LEAD CONTACT WEBSITE

Dachuan Zhang: https://www.fst.nus.edu.sg/our_people/faculty-members/zhang-dachuan/